m

Journal of Experimental Psychology

150.57
B2

ARTHUR W. MELTON, Editor
HUMAN RESOURCES RESEARCH CENTER
LACKLAND AIR FORCE BASE
SAN ANTONIO, TEXAS

CONSULTING EDITORS

Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Robert M. Gagné, *Human Resources Research Center*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
David A. Grant, *University of Wisconsin*
Lloyd G. Humphreys, *Human Resources Research Center*
Donald B. Lindsley, *University of California, Los Angeles*
Neal E. Miller, *Yale University*
Edwin B. Newman, *Harvard University*
Kenneth W. Spence, *State University of Iowa*
Benton J. Underwood, *Northwestern University*
Delos D. Wickens, *Ohio State University*

Lorraine Bouthilet, Managing Editor

THE JOHN CRERAR LIBRARY
RECEIVED
AUG 4 - 1953

V.46 - Jan

CONTENTS

The Discrimination of Dot Patterns as a Function of Number and Average Separation of Dots: R. S. FRENCH 1

Previous Training as a Determinant of Response Dependency at the Threshold: W. S. VERPLANCK, J. W. COTTON, AND G. H. COLLIER 10

The Psychophysiological Significance of the Galvanic Skin Response: A. C. MUNDY-CASTLE AND B. L. MCKIEVER 15

Physiological Need, Word Frequency, and Visual Duration Thresholds: L. G. WISPÉ AND N. C. DRAMBAREAN 25

An Electromyographic Study of Tension in Interrupted and Completed Tasks: A. A. SMITH 32

The Effect of Difficulty of Task on Proactive Facilitation and Interference: A. A. BARCH 37

Acquisition of Two Lever-Positioning Responses Practiced over Several Periods of Alternation: E. A. BILODEAU 43

Some Effects of Manifest Anxiety on Mental Set: I. MALTZMAN, J. FOX, AND L. MORRISETT, JR. 50

Anxiety, Anxiety Reduction, and Stress in Learning: J. DEESE, R. S. LAZARUS, AND J. KEENAN 55

Negative Transfer in Verbal Learning: L. W. PORTER AND C. P. DUNCAN 61

American Psychological Association

Vol. 46 No. 1 July 1953

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc. The subscription rate per volume is $7.50, or $15.00 annually. Single copies are $1.50. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology. Normally, articles of a length exceeding 15 printed pages cannot be accepted. However, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) must be presented in a single article, rather than in a series of articles.

Articles are published in the order of their receipt, except in rare circumstances. Authors who request them are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas. Priority in publication is given to articles whose authors assume the full cost of publication.

Address all articles submitted for regular or priority publication to the editor, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 (4, Suppl.), 389–449. In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Journal of
Experimental Psychology

VOL. 46, NO. 1 JULY, 1953

THE DISCRIMINATION OF DOT PATTERNS AS A FUNCTION OF NUMBER AND AVERAGE SEPARATION OF DOTS [1]

ROBERT STANTON FRENCH

Perceptual and Motor Skills Research Laboratory, Human Resources Research Center

The present study investigates the task of discriminating differences between dot patterns. The problem for *S* was one of indicating whether pairs of patterns were the "same" or "different." When a "difference" appeared, it consisted of a small, but constant, displacement of one of the dots in one of the two patterns to be compared. It appeared reasonable on a priori grounds that as the number of dots in a pattern is increased, the discrimination of the slight difference should become more difficult, since the proportion of the pattern not changed is increased. On the other hand, a pattern might be composed of so few dots that the discrimination of a difference is primarily a discrimination of length, or of slope. Accordingly, it was hypothesized that between two-dot patterns and possibly seven-dot patterns there should be a number which produced optimal discriminations of pattern difference.

The choice of the average separation of the dots as a variable to be manipulated was based on the expectation that the displacement of a dot would be more readily detected if the amount of the displacement were large relative to the average separation of the dots as a group. Accordingly, it was hypothesized that the greater the average separation of the dots in a pattern, the more difficult would be the discrimination of the small constant displacement of one of the dots.

A search of the literature revealed few similar attacks on this problem of pattern discrimination. However, a considerable number of recent research studies have dealt with the general problem of the discrimination of visual number, e.g., Hunter and Sigler (**2**), Taves (**7**), Saltzman and Garner (**6**), Kaufman, Lord, Reese, and Volkmann (**3**), and Casperson and Schlosberg (**1**). Typically, the task is that of viewing a collection of dots or similar elements and indicating the number present. Although there are a number of superficial resemblances between the present study and the studies on visual number cited, a fundamental difference remains. In the present study, a

[1] The experimental work for this study was performed as part of the United States Air Force Human Resources Research and Development Program. The opinions or conclusions contained in this report are those of the author. They are not to be construed as reflecting the views or indorsement of the Department of the Air Force.

given *S* is never asked to discriminate between patterns differing in *number* of dots, i.e., for a given *S* the number of dots represented a constant or parameter of the experimental situation. The number of dots in a pattern was studied as an independent variable, but only as it might affect discrimination within a given dot condition.

Leyzorek (**4**) investigated the discrimination of the spatial displacement of a point of light as a function of the temporal interval between the presentation of the light point in two positions. Perhaps of greatest interest to the present discussion was the finding that the single fixation point in the darkened room was as effective for good spatial discrimination as the same fixation point in a fully lighted room. Leyzorek suggested that the function relating spatial discrimination to visual-field complexity may not be monotonic. The present study would seem to offer a test of this hypothesis to the extent that visual-field complexity may be defined in terms of the number of dots in a pattern.

A group of studies by Nixon (**5**) utilized patterns of dots in a task of reconstruction from immediate memory. Over a series of experiments the number of dots was varied from one to five and the time between exposure and reconstruction from 3 to 10 min. Nixon found that the time interval, the spatial position of the dots relative to a circular frame, and the number of dots were significant sources of variance. When the number of dots in the pattern was increased from one to three, a progressive increase in the average error was noted. The four- and five-dot patterns produced no further increase in error, however. Nixon interpreted this result, with some reservations, as contrary to the Gestalt notion of greater structural stability and consequently better retention for the more complex patterns.

Method

Apparatus.—This study utilized a device designated as the Pattern Discrimination Apparatus. By means of the apparatus series of patterns varying widely in number of dots and specific configuration were presented to groups of eight *Ss* each. To the *Ss*, seated about 15 ft. from the apparatus in a darkened room, the patterns appeared as pin points of light against a homogeneous background. A plywood screen about 6 ft. long and 4 ft. high constituted the background and served the purpose of hiding from *Ss* all details of operation of the apparatus. The patterns of light points appeared on the screen by means of light projected from small individual light bulbs through pin holes in a metal plate to a large sheet of translucent white paper which covered the entire screen. An almost limitless variety of patterns could be obtained since 112 such light sources were arranged in rows and columns 1 in. apart forming a matrix of circular outline. The apparatus consisted of two such circular matrices of light sources centered on the screen side by side, approximately 22 in. from center to center. Each of these matrices of bulbs was mounted in a circular frame about 14 in. in diameter. Two circular holes were cut in the screen and the two frames were fitted into these holes in such a way that the metal plate containing the pin holes was flush against the back side of the white paper covering the screen. Each of these circular frames was an independent unit held in place by roller bearings and free to be rotated 180° either to the right or left, thus rotating the pattern as seen by the *S*. In this particular study, however, the rotation feature of the apparatus was not utilized. Each bulb in each of the circular matrices was wired individually to a corresponding switch in one of two banks of such switches.

The apparatus was designed to be used for a number of studies all dealing with the general problem of the perception of dot patterns but differing in specific detail. By having two independent ensembles of possible points, two patterns could be presented either simultaneously or with any desired exposure interval between. Additional controls determined whether the pattern on the right or left side of the screen appeared first. If "knowledge of results" was to be investigated, a printed "S" or "D" indicating "same" or "different" could be illuminated automatically after the presentation of a pair of patterns; at the same time the pair of patterns would reappear permitting a careful examination of the forms. To summarize the kinds of presentation with the

apparatus: any pattern of light points or pair of patterns desired could be set up by throwing the appropriate switches in the two banks. With another switch the apparatus could be set for either simultaneous or successive presentation of the two patterns. By rotating one of the circular matrices, the two patterns could be presented with different spatial orientation. If the presentation was to be successive, the setting of another switch determined whether the pattern on the right or left matrix appeared first. Three electronic interval-timers could be set to control the duration of exposure for each pattern of a pair and the interval between; a fourth timer controlled the time elapsed before the patterns reappeared together with the illumination of the "S" or "D" if this was desired; a fifth timer controlled the duration of this representation.

With the controls set appropriately for the presentation of a specific pattern or pair of patterns, *E* threw a single switch which caused some combination of the events described above to occur automatically in sequence.

Each *S* had a covered box on the arm of his chair containing two response keys appropriately labeled. Both the kind of response and its latency measured to .01 sec. could be obtained. Eight electric chronoscopes were wired into the circuit to start either at the time of presentation of the two patterns, if simultaneous, or at the time of the presentation of the second pattern of the pair, if successive. The response made by *S* was indicated by two colored lights, one of which was illuminated at the time *S* pressed one of the response keys. The response and its latency were recorded manually.

The room was dimly lit from lamps behind the screen. These lights illuminated the controls and effectively darkened the front of the screen by accentuating the contrast between the screen and its immediate surroundings. The perceptual task for the *S*s viewing patterns with the apparatus utilized essentially only foveal vision, the visual angle subtended by an average pattern being about 1° 26′.

Patterns.—For the present study, 120 dot patterns were designed for each number of dots from two to seven in accordance with the following principles: For each group of patterns the average separation of the dots was varied systematically within the limits of the apparatus; the dots were spaced more evenly than they would have been if randomly determined; vertical, horizontal, or regular angular arrays of dots, etc. were avoided. At the same time, care was taken to make the patterns in a group as varied as possible. The only difference introduced between the two patterns to be judged "same" or "different" was a 1-in. displacement (a visual angle of approximately 18′) of one of the dots. The spatial position of this dot was varied as much as possible. For each group of *S*s half of the 120 trials were "same" and half "different" in a nonsystematic sequence. Figure 1 presents examples of the magnitude of the difference to be judged. Single examples of pairs of patterns from the two-, three-, and seven-dot series are illustrated. First Pattern A was exposed 3 sec. and then Pattern B appeared for 3 sec.

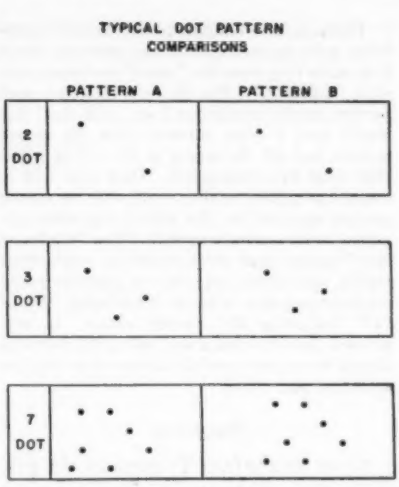

FIG. 1. Examples of pairs of patterns from the two-, three- and seven-dot series illustrating the magnitude of the difference to be judged

Each of the 120 patterns was presented once in a sequence constructed using random numbers. Within consecutive blocks of ten trials the time between the presentation of the standard and comparison stimulus was varied nonsystematically from 1 to 10 sec. Thus, each of ten time intervals was tested individually with 12 different patterns.

To insure that each pattern was tested both as an "S" and a "D" pattern, opposite forms of the procedure were employed. Half of the 32 *S*s within each dot condition were tested with the specific designation of "S" and "D" trials reversed relative to the other half of the *S*s. With respect to number of dots and "knowledge of results" (hereafter called "reinforcement") the experimental design was factorial. Each of the six dot conditions was tested both with and without reinforcement.

Subjects.—The *S*s were 384 basic airmen trainees at Lackland Air Force Base; 32 *S*s were assigned nonsystematically to each of 12 major conditions.

Instructions.—The instructions described carefully, with examples, how the patterns would look when they were the "same" and when they were "different." The *S*s were told that each pattern would appear for 3 sec. and that they would have a 6-sec. interval after the second pattern had left the screen at the end of which they must have responded. They were told to respond as quickly as they could after the second pattern appeared on the screen, but were cautioned to be accurate as well. The "reinforcement" groups were given additional instructions stating that after 6 sec. the two patterns would reappear together with an illuminated "S" or "D" indicating the correct choice. It was stressed that if a choice was wrong the patterns should be checked carefully to determine why an error had been made.

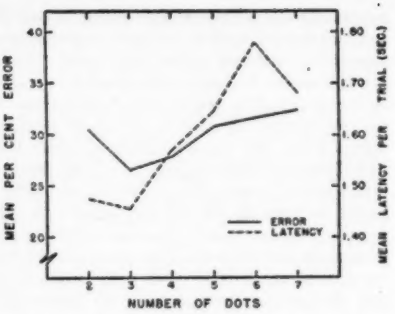

FIG. 2. Mean per cent error and mean latency per trial as a function of number of dots for the two reinforcement groups combined ($N = 64$ each point)

RESULTS

Error analysis.—To present the primary data of the study, the total error and mean latency was determined for each *S* tested with each dot condition. The error data are shown in Table 1 and Fig. 2. Both the means and *SD*'s are expressed in terms of percentage error. In Fig. 2 the error and latency data have been plotted on the same graph for easier comparison. For each measure, individually, the data for the reinforcement and no-reinforcement groups have been combined.

It is evident from Fig. 2 that the number of dots in the pattern systematically affects accuracy of discrimination. An analysis of variance of the error data revealed that the difference between number of dots was highly significant ($p < .001$; $F = 11.85$; $df = 5/372$), and the difference between the reinforcement conditions was significant at the .05 level ($F = 5.12$; $df = 1/372$). The interaction between these variables was not significant ($F = 1.16$; $df = 5/372$). *t*-test comparisons were made between groups representing successive increases in number of dots and between the three- and seven-dot groups. The mean square within cells from the analysis of variance was used to derive an error term. In brief, the differences between the means of the two- and three-dot groups and between the three- and the seven-dot groups were all found to

TABLE 1

MEAN PER CENT ERROR AND *SD*'S FOR EACH NUMBER OF DOTS WITHIN THE REINFORCEMENT (R) AND NO-REINFORCEMENT (No-R) CONDITIONS
(N in each group = 32)

Reinforcement Condition	Number of Dots											
	2		3		4		5		6		7	
	Mean %	*SD*	Mean %	*SD*	Mean %	*SD*	Mean %	*SD*	Mean %	*SD*	Mean %	*SD*
R	29.43	6.34	25.55	5.50	26.82	4.68	30.81	4.65	30.62	5.57	33.02	5.93
No-R	31.54	5.11	27.55	5.69	28.91	4.79	30.96	4.13	32.89	5.68	31.80	4.54
Total	30.48		26.55		27.86		30.88		31.76		32.41	

TABLE 2

LATENCY MEANS AND *SD*'S (IN SECONDS) FOR EACH NUMBER OF DOTS WITHIN THE REINFORCEMENT (R) AND NO-REINFORCEMENT (No-R) CONDITIONS
(*N* in each group = 32)

Reinforcement Condition	Number of Dots											
	2		3		4		5		6		7	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
R	1.46	0.24	1.49	0.30	1.61	0.26	1.72	0.30	1.77	0.24	1.61	0.23
No-R	1.49	0.35	1.42	0.29	1.53	0.26	1.57	0.32	1.78	0.28	1.75	0.27
Total	1.477		1.456		1.570		1.643		1.779		1.681	

be significant at the .01 level. (two-tailed test; $df = 372$). It may be stated with considerable confidence, therefore, that the greater accuracy of discrimination with three-dot patterns is not a chance occurrence.

Latency analysis.—Similar analyses were performed for the latency data. Table 2 presents the mean latency per trial for each group of Ss together with the *SD*'s of the corresponding distributions. The latency data with the reinforcement and no-reinforcement groups combined are plotted in Fig. 2 together with the error data. An analysis of variance for the latency data revealed that number of dots was again a significant effect at the .001 level ($F = 12.12$; $df = 5/372$). Neither the difference between the reinforcement conditions ($F = .05$; $df = 1/372$) nor the interaction of reinforcement and number of dots was significant ($F = 2.07$; $df = 5/372$). For the reinforcement groups as determined by the t test, deriving an error estimate from the analysis of variance, the difference between the means of the six- and seven-dot groups was significant at the .01 level (two-tailed test; $df = 372$). Since both the reinforcement and no-reinforcement groups show a sharp drop in latency at this point, there is reason to believe that the effect is a true one.

Learning data.—Figure 3 is a learning function for all of the dot conditions and the reinforcement conditions combined. Mean percentage of error has been plotted for 12 successive blocks of ten trials each. Both the errors on "S" and "D" trials and the total errors are shown. It may be seen that some learning occurs as the experiment progresses. On the face of it, however, the "D" accuracy appears to get worse after the seventh block of trials, while "S" accuracy continues to improve. Note also that the "S" and "D" functions after the seventh block are quite obviously inversely related, while the total error function remains relatively flat. A possible explanation for the rather puzzling effects illustrated in Fig. 3 will be discussed later.

Dot separation.—It was believed likely that the average separation of the dots within each pattern was an important source of variability in performance from pattern to pattern. The procedure of analyzing the effect of this variable involved measuring the distance from the displaced dot to each other dot in the pattern. This was done for each of the dot patterns used. Then the logarithms of these distances were found and the mean logarithm value for each pattern was determined. Essentially linear scatter plots were obtained when these mean

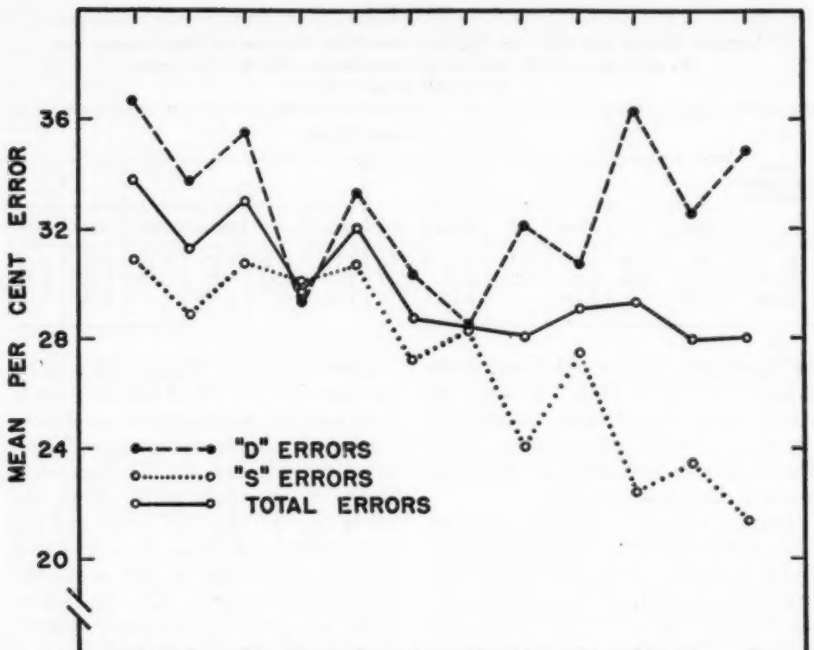

FIG. 3. A learning function for all of the dot conditions and reinforcement conditions combined. Mean per cent error for the "S" and "D" errors separately and for total error is plotted for 12 consecutive blocks of ten trials each.

logarithm values were plotted against the errors made on the corresponding patterns (thus justifying the use of a logarithmic conversion). "S" and "D" errors were plotted separately, with the reinforcement groups combined. These coefficients for each of the dot conditions are shown in Table 3. It will be observed that all of the "D" coefficients are significantly different from zero at the .001 level of confidence, ranging from .55 to .79. With the "S" coefficients ranging from .18 to .32, only two achieve this .001 level. In other words, Table 3 indicates that when the patterns were "different," the average separation of the dots influenced accuracy to a

TABLE 3

COEFFICIENTS OF CORRELATION BETWEEN MEAN LOG INTER-DOT DISTANCE FOR EACH PATTERN AND "S" AND "D" ERRORS SEPARATELY
(N for each correlation = 120)

Type of Error	Number of Dots					
	2	3	4	5	6	7
"D"	.567	.766	.663	.732	.546	.789
"S"	.322	.291	.223	.324	.179	.183

* For significance at the .05, .01, and .001 levels r must be .18, .23, and .30, respectively.

highly significant extent and to a considerable, but lesser, extent when the patterns were "same." Note that increasing the number of dots in a pattern necessarily decreases their separation on the average, if the patterns are within a circumscribed area as in the present study. We computed the mean separation for each group of patterns and found that as the number of dots increased, their average separation did, in fact, decrease. However, since accuracy of discrimination becomes progressively *poorer* beyond three dots in spite of this decreased separation, some factor other than average separation must come into the picture.

Discussion

A restatement of the main findings may be helpful at this point: (*a*) three-dot patterns were found to yield the most accurate discrimination of a small difference; (*b*) the time required to respond increased steadily up to six dots and then dropped sharply; (*c*) in general, giving *S* "knowledge of results" after each judgment produced a small but consistent increase in accuracy; (*d*) both the "S" and "D" errors were found to correlate positively with a measure of the average separation of the dots in each pattern. Considerably higher correlations were found for the "D" errors, however.[2]

The results indicate that an individual is better able to detect a slight difference in the relative positions of three dots than of either two dots or any greater number up to seven. It does not necessarily follow, however, that three-dot patterns are most easily *differentiated*. In the present task *S* compares each pattern with another almost identical to it. The task of differentiation, on the other hand, would require learning a different response for each of a number of patterns. These two procedures might yield quite different functions.

It is fairly evident that a number of variables are involved in the relation of accuracy to number of dots. Tentatively, one of these variables might be identified as the number of relations which are modified by the displacement of one dot. With only two dots in a pattern an individual may do little more than detect a difference separating the two dots, or a difference in both the distance and the slope of an imagined line connecting the two dots. With three dots there are three such distance relations which may serve as cues, as many slopes, and innumerable possible "articulations" of the visual field in terms of angles, bisections, vertical and horizontal extrapolations, etc. Clearly, as the number of dots increases, the number of relations modified by a displacement increases as well. On the other hand, the probability of utilizing these relations to discriminate a difference appears to decrease beyond three dots. As a possible explanation, note that although the absolute number of modified relations increases, the relative proportion actually decreases. Consider, for example, the number of distance relations among three dots. When one dot is displaced a slight amount, two out of the three distances are changed. With four dots there are six possible relations of pairs of dots. With one of these displaced, three out of six, or one-half, of the relations are modified.[3] Since the equation expressing this proportion describes a negatively accelerated function, its form would be correct to account for the shape of the error function from three dots to seven. The drop in errors from two to three dots, however, cannot be accounted for on this basis. Possibly a qualitative change in the kind of relations available occurs between two and three dots.

The increase in the time required to respond (Fig. 2) as the number of dots increases from two

[2] It should be remembered that an "S" error is actually a "D" response on an "S" trial and vice versa. The kind of error was defined by the physical identity or dissimilarity of the patterns compared on that trial.

[3] A general equation expressing this proportion may be written as follows:

$$\frac{R_m}{R_t} = \frac{n-1}{\frac{n(n-1)}{2}}$$

where R = the number of distance relations among dots, m = modified relations, t = total relations, and n = number of dots. This equation may be simplified to:

$$\frac{R_m}{R_t} = \frac{2}{n}$$

Thus, twice the reciprocal of n yields the proportion of distance relations modified by the displacement of *one* out of any given number of dots.

to six is an expected result. On the other hand, the sharp drop in latency from six to seven dots requires comment. A number of recent studies have investigated the discrimination of visual number (1, 2, 6, 7). Kaufman, Lord, Reese, and Volkmann (3) have rather consistently found that *S*s respond differently to numbers of dots up to six as compared with six or more. They find this phenomenon with accuracy, report time, and a measure of confidence. Although the present task is different, it is rather interesting to find that they report an increase in the time required to report as the number of objects increases from one to six, but little further increase beyond six dots. One might account for an actual drop in latency, as in the present study, in terms of an abrupt change in the characteristics of the task. One might expect that *S* could retain, over a short interval, a rather complete and accurate impression of patterns up to six dots, but perhaps not beyond. For patterns of not more than six dots, *S* might be expected to make a rather methodical "check" of corresponding points of the present stimulus and his retained "impression," and then abandon this method for patterns of greater number.

The finding that the average distance separating the dots is an important factor for discrimination accuracy has a number of implications. By squaring the correlation coefficients presented in Table 3, it may be seen that for the two-dot condition average separation of the dots in the pattern accounted for approximately 32% of the total variance in "D" errors and that for the seven-dot condition this factor accounted for about 61% of the total variance. The surprising aspect of this result is that the variance predictable from average separation is not considerably reduced in the more complex patterns. One might expect various "grouping" responses, associations with familiar objects, etc. to become increasingly important as the number of dots increased. This would not appear to be the case, at least in a situation where the difference to be detected is only one part of the pattern (i.e., one dot). Note that the difference to be detected was a constant amount regardless of the average separation of the dots in the pattern. Accordingly, the results of the present study are only indirectly applicable to situations in which the scale of the total pattern, including areas of deviation, is modified.

With reference to Fig. 3, which presented a learning function for the "S" and "D" error curves separately, it has been pointed out that on the face of it, "D" accuracy appears to grow worse after the seventh block of trials, while "S" accuracy continues to improve. Seeking a parsimonious explanation for this effect, certain mathematical relationships become evident. It may be shown by relatively simple algebra that the difference between the number of "S" and "D" *errors* in the present two-choice situation bears a linear relationship to the relative proportion of "S" and "D" *responses*.[4] Accordingly, if one knows the proportion of the total number of responses which are "S" and the proportion which are "D," a little simple arithmetic will yield the error difference which must necessarily be present. With this relationship taken into consideration, the separation of the "S" and "D" error functions in Fig. 3 takes on new meaning. It is evident that for approximately the first half of the testing period the proportion of each kind of response made by the *S*s reflects fairly accurately the proportion of trials of each kind in the test series. From the seventh block on, however, there is no further reduction in total error, only an increase in the number of "S" responses made. An explanation for this effect suggests itself: perhaps some aspect of the training procedure alters the individual's difference threshold such that an increasingly large difference is needed to evoke the response "D." In such a two-choice situation this would have the effect of increasing the number of "S" responses. Possibly this change in threshold results from accumulating work inhibition or "fatigue." It is less likely that learning is involved, since the effect of repeated examples of the difference to be detected should make *S* more familiar with the amount and nature of this difference. Any condition producing a relative "stimulus impoverishment," such as a very short exposure time, a long retention interval, certain kinds of distraction, etc., might be expected to produce a comparable increase in "S" responses.

Regardless of the true explanation for such a change in the proportion of "S" and "D" responses, the present data suggest that caution should be exercised when one attempts to interpret performance curves in a two-choice situation. One need not necessarily conclude from these data, for example, that the judgment of identity and difference are "qualitatively" different tasks. It is suggested that any condition producing a relative "stimulus impoverishment" will lead to an increase in "S" responses, and thus, in the two-choice situation, a necessary separation of the two kinds of error curves.

Summary and Conclusions

1. The discrimination of differences between dot patterns was studied as a

[4] $\frac{\text{"S" responses} - \text{"D" responses}}{2} =$ "D" errors − "S" errors.

function of the number of dots and their average separation.

2. The Pattern Discrimination Apparatus was used to present a series of 120 different patterns for each number of dots from two to seven. Each pattern appeared twice, the pair of patterns being either identical or differing by the slight displacement of one of the dots. The *Ss* responded "Same" or "Different." The *Ss* were 384 basic airmen trainees at Lackland Air Force Base.

3. The results indicate that: (*a*) *Ss* are better able to detect a slight difference in the relative positions of three dots than of either two dots or any greater number up through seven; (*b*) the average separation of the dots is a highly significant source of variance; the more widely spaced the dots in the pattern, the less accurate is the discrimination of a constant displacement; (*c*) when knowledge of results is given immediately after each judgment, *Ss* in general achieve a higher level of accuracy.

4. It is suggested, as a possible explanation for the decrease in accuracy beyond three dots, that an individual responds to the proportion of relations modified by the displacement of a dot rather than the absolute number of such relations. Certain simple relationships between responses and errors in a two-choice situation are discussed in connection with the concept of "stimulus impoverishment."

References

1. Casperson, R. C., & Schlosberg, H. Monocular and binocular intensity thresholds for fields containing 1–7 dots. *J. exp. Psychol.*, 1950, **40**, 81–92.
2. Hunter, W. S., & Sigler, M. The span of visual discrimination as a function of time and intensity of stimulation. *J. exp. Psychol.*, 1940, **26**, 160–179.
3. Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. The discrimination of visual number. *Amer. J. Psychol.*, 1949, **62**, 498–525.
4. Leyzorek, M. Two-point discrimination in visual space as a function of the temporal interval between the stimuli. *J. exp. Psychol.*, 1951, **41**, 364–475.
5. Nixon, S. R. *Some experiments on immediate memory.* Medical Research Council, Applied Psychology Research Unit, Psychological Laboratory, Cambridge, England. Report No. A.P.U. 39, November, 1948.
6. Saltzman, I. J., & Garner, W. R. Reaction-time as a measure of span of attention. *J. Psychol.*, 1948, **25**, 227–241.
7. Taves, E. H. Two mechanisms for the perception of visual numerousness. *Arch. Psychol.*, 1941, **37**, Whole No. 265.

(Received November 17, 1952)

Journal of Experimental Psychology
Vol 46, No. 1, 1953

PREVIOUS TRAINING AS A DETERMINANT OF RESPONSE DEPENDENCY AT THE THRESHOLD [1]

WILLIAM S. VERPLANCK,[2] JOHN W. COTTON,[3] AND GEORGE H. COLLIER [4]

Indiana University

In a recent paper (4) the authors presented the results of an investigation of the statistical properties of responses to repeated visual stimulations. On each of four days, 16 human *Ss* were stimulated 300 successive times by a patch of light of fixed duration, dimensions, position, and brightness. The data were analyzed for the statistical independence of successive responses by means of a serial-correlation technique, and sequential dependencies were found. Hence, assumptions that successive responses are randomly arranged and statistically independent (often implicit in psychophysical measurements and in theories of the visual threshold) cannot be made for any specific situation in the absence of a strict empirical test. The probabilities of a "yes" following a "yes" and of a "no" following a "no" are greater than would be expected from the values estimated from the over-all relative frequencies of "yeses" or "noes" in the whole series of trials.

In the above experiment *Ss* were given five days of preliminary training on a method-of-limits procedure that presents stimuli in intensity-ordered series. A brief method-of-limits session also preceded and followed the single-brightness series on the four experimental days. This procedure necessarily leads *Ss* to give long groups or sequences of "yeses" to stimuli of the higher intensities followed by similar groups or sequences of "no" responses to those of lower intensity. Some might infer that such repetitions of "yes" and "no" in series produce a tendency to give similar sequences of "yeses" and "noes" when all stimuli are of the same brightness.

To test this possibility, we have employed two methods of threshold determination in the preliminary training and in the short sessions which precede and follow single-brightness stimulations. The first method is the same as that used previously; it necessarily produces sequences of "yeses" and "noes." The second method employs a series of stimuli randomly varied in brightness so that few, if any, long sequences of the same responses may be expected to occur. Evidence of nonindependence is then sought in the responses to 300 single-brightness stimulations of two groups of *Ss*, of which one has been trained according to the first procedure, and the other according to the second.

It may be argued that a "set" for, or "expectancy" of, stimuli of graded intensity will be set up in members of the first group and hence that they will tend to repeat responses and so exhibit serial dependencies. No such expectancy should be set up by the training procedure in members of the second group, and hence they should not exhibit serial dependencies.

[1] The experiment reported in this paper was performed in the Department of Psychology, Indiana University, under ONR Contract N6onr-180, T. O. IV, Project 143–253, and with support from a grant by the Graduate School of Arts and Sciences. Work is continuing at Harvard University under ONR Contract N5ori-07639.

[2] Now at Harvard University.

[3] Now at Northwestern University.

[4] Now at Duke University.

Apparatus and Procedure

Apparatus.—The apparatus used in this experiment was the Model III Hecht-Schlaer Adaptometer (2), modified as described in a preceding paper (4).

Subjects.—Six paid Indiana University male students, two graduates and four undergraduates, served as *Ss*. All met or exceeded Service visual standards as tested by the Orthorater. All were previously inexperienced in visual experimentation and were assigned at random to one or the other of the two experimental groups.

Dark adaptation and instructions.—The *Ss* were run fully dark-adapted. The dark-adaptation procedure and the instructions in this experiment were the same as those already described (4).

Procedure.—The experimental procedure may be outlined as follows: A one-day session of threshold determination served as a training period and preceded four days in which *Ss* of each group were given repeated single-brightness stimulations in order to test for randomness of response. The treatments of two groups of three *Ss* each differed in the methods of stimulation followed during the training session and in the procedures that preceded and followed the single-brightness series on the other four days.

In the preliminary session (Day A) members of Group I received 20 series of intensity-ordered stimulations as described elsewhere (5). These gave each *S* approximately 225–250 stimulations. The members of Group II received 240 randomized stimulations with brightnesses ranging from 1.9 to 3.0 log $\mu\mu$L in .1 log-unit steps. The value of the 50% threshold obtained for each *S* on this first day was used as the single brightness employed in subsequent sessions. These values are presented in Table 1.

On Days 1 and 2 the members of Group I received three pairs of descending and ascending series of intensity-ordered stimulations, followed by 300 stimulations at the 50% threshold. The experimental day ended with six more series of intensity-ordered stimulation. On Days 1 and 2, members of Group II received 60 stimulations with randomized brightnesses before 300 single-brightness stimulations at their 50% thresholds. Another 60 randomized stimulations followed the single-brightness series. A comparison of the incidence of sequences on Days 1 and 2 for the two groups should indicate the relative effects of the two training procedures.

The procedures followed on Days 3 and 4 were exactly like those on Days 1 and 2, except that the treatments of the two groups were interchanged. Group I received the random series of brightnesses before and after the single-brightness tests, and Group II received the six intensity-ordered series before and after single-brightness tests.

Results

Table 1 presents the relative frequencies of "yes" for each series of single-brightness stimulations. The mean relative frequency (P_R) of "yes" is .49, and the *SD* of these is .19. These values are not significantly different from those reported earlier in a larger group of *Ss* (mean relative

TABLE 1

50% Thresholds, Relative Frequencies of "Yes" (P_R), and CR_R's by Subjects and Days

(300 presentations in each cell)

Group	*S*	Initial 50% Threshold (log $\mu\mu$L)	Day 1		Day 2		Day 3		Day 4	
			P_R	CR_R	P_R	CR_R	P_R	CR_R	P_R	CR_R
		(limits)	(limits)		(limits)		(random)		(random)	
I	1	2.48	.12	2.44	.26	8.39	.41	9.03	.39	5.28
	2	2.64	.79	.52	.64	7.36	.57	4.45	.75	2.60
	3	2.20	.24	2.04	.29	2.06	.55	3.36	.47	1.15
		(random)	(random)		(random)		(limits)		(limits)	
II	4	2.60	.53	3.70	.41	2.58	.28	2.38	.35	3.68
	5	2.42	.82	.56	.74	.18	.84	2.76	.52	5.58
	6	2.50	.42	6.03	.51	5.45	.41	4.01	.41	3.54

frequency = .55; SD = .25) and, as before, are not significantly different from the .50 predicted from the first day's results.

The analysis of results to determine whether or not the responses within the blocks of single-brightness stimulations are statistically independent of one another was performed by the test for serial correlations described by Hoel (3). The use of this method has been discussed elsewhere (4). Table 1 presents the critical ratios (CR_R) of the serial correlation coefficient for each series of single-brightness stimulations. Of the six CR_R's of the *Ss* of Group I on Days 1 and 2, three exceeded the 5% level of significance and two exceeded the 1% level. Of the six comparable tests for Group II, four exceeded the 1% level and two failed to reach the 5% level. Both groups demonstrated nonrandom arrangements during Days 1 and 2, although Group II had received training calculated to minimize the possibility of development of a response set towards repetition of responses from trial to trial. On Days 3 and 4 when the antecedent stimulation procedures are reversed, five of the six serial-correlation values of Group I are significant at the 1% level, and Group II also shows five coefficients exceeding the 1% level, with the other exceeding the 5% level. The groups do not differ, but there is the suggestion of a trend towards increased serial correlations with successive days.

The possibility of changes within a group attendant upon the change of procedure beginning on Day 3 must be considered. The number of significant correlations found in each of the two-day periods remained constant for members of Group I. Group II shows an increase from four significant values on Days 1 and 2 to six on Days 3 and 4. These findings suggest that a change from intensity-ordered to random training procedures in Group I did not produce independence of successive responses while a change from random to intensity-ordered training procedures may have increased the degree of dependence of successive responses. Thus the hypotheses with which we are concerned may seem partially verified; although an appropriate history of responses is not at all necessary for the appearance of the sequential effect, it may increase it in magnitude. A further analysis was performed.

The experimental design employed is such that a crossover analysis of variance (1) may be made on the data. This analysis closely resembles the latin square, and is suitable to the present data since our number of treatments is two. A test of training effects was therefore performed using a modified crossover-design analysis of variance. The data were first grouped into blocks of two days each. Following the crossover-design method, the data were then broken down into sums of squares for subjects, blocks of days, treatments, and error. A further breakdown of the sum of squares for *Ss* into a sum of squares for groups (differing in the order of treatments) and a sum of squares for *Ss* within groups was made. Thus the possibilities that training and the history of training affected the data could both be tested. Analysis-of-variance tables based on these sums of squares are presented in Table 2 for both the relative frequencies and the CR's. Only the differences between *Ss*' frequencies of response proved significant. Groups, treatments, and blocks of days apparently have no effect upon either the relative frequency or the CR's. One qualification must be stated. In our experimental design, the training procedure

TABLE 2

ANALYSES OF VARIANCE OF RELATIVE FREQUENCY OF "YES" RESPONSES (F_{P_R}) AND OF C_{R_P}'s

Source	*df*	Mean Square for Frequency of Response		*p*	Mean Square for Critical Ratios		*p*
		V_{P_R}	F_{P_R}		V_{CR_R}	FCR_R	
Groups	1, 4	4,256	<1.00	>.05	6.1490	<1.00	>.05
Subjects within groups	4, 4	26,942	*10.11*	<.05	24.2327	2.31	>.05
Blocks of days	1, 4	243	<1.00	>.05	3.9331	<1.00	>.05
Treatments	1, 4	15,265	5.73	>.05	.0469	<1.00	>.05
Error	—	2,666	—		10.4828	—	

during the preliminary sessions and the temporal order of treatments are combined into one estimate of variance, that for groups. If both variables were effective, but operated in opposite directions, the same results might be obtained. We must conclude that either the confounded variables have no effect or that their effects are equal and opposite. The latter is unlikely.

A possibility that must be considered is that nonindependence of responses may arise as a result of slow temporal trends in the relative frequency of response (P_R). If, for example, *S* gives almost all "yeses" during the first portion of the session and almost all "noes" during the second portion, a significant CR_R will result. This follows from the facts that the serial correlation test is sensitive to trend, as well as to statistical nonindependence of successive responses, and that it does not distinguish between the two effects.

This possibility was tested by the use of Pearsonian correlation. The number of "yeses" to each successive set of ten stimulations was correlated with the ordinal position of the set in the series of 300. Three *S*s yielded low (about .30) negative correlations that met the 5% level of significance, between P_R and ordinal position; a small decrease in P_R occurred throughout the experimental period. The other three *S*s showed near-zero correlations that were not significant. However, highly significant sequential effects occurred not only in the data of *S*s showing this decrement, but also in those of the *S*s who did not.[5] A trend in time may be operating, but it is not adequate to account for the nonrandomness observed.

DISCUSSION

The results of this experiment lend little credence to the hypothesis that the lack of randomness found in responses to invariant visual stimuli in our previous experiment (4) can be attributed to the training procedures used in that experiment. In the present study, nonrandomness characterized by *long* series of similar responses ("yeses" or "noes") was obtained following training either by randomly ordered or intensity-ordered stimuli.

A second factor, that of trends within sessions, may be present in our experimental data. Again, the absence of sizable correlations between the number of "yeses" per group and the ordinal position of the group, shows that such trends cannot account for

[5] In other, as yet unpublished, data we have evaluated CR_R on blocks of 50 trials—a set short enough so that no trend can be observed; CR_R's are still significant statistically, although less so than in the present data, as they must be owing to the decreased sample size.

the nonrandomness found. This nonrandomness is evidence for nonindependence of successive responses.

Summary

This experiment consisted of five experimental sessions, one devoted to an original threshold determination for each of six *S*s and the other four devoted to 300 successive single-brightness stimulation periods per session. The six *S*s were divided into two groups of three *S*s each. Group I was given intensity-ordered stimulations in all threshold determinations on the first three days of the experiment; Group II was given randomized stimulus brightnesses at the comparable times. On the last two days of the experiment the procedures for Groups I and II were reversed.

Despite the differences in training procedures for the two groups, both groups exhibited nonindependence of responses in the single-brightness series.

Some evidence was found for a linear trend toward decreased frequency of seeing late in each experimental session. Although this factor alone might produce nonrandomness of response, there is strong evidence that other mechanisms are operative.

References

1. Cochran, W. G., & Cox, G. M. *Experimental designs.* New York: Wiley, 1950.
2. Hecht, S., & Shlaer, S. An adaptometer for measuring human dark adaptation. *J. opt. Soc. Amer.*, 1938, **28**, 269–275.
3. Hoel, P. G. *Introduction to mathematical statistics.* New York: Wiley, 1947.
4. Verplanck, W. S., Collier, G. H., & Cotton, J. W. Nonindependence of successive responses in measurements of the visual threshold. *J. exp. Psychol.*, 1952, **42**, 273–282.
5. Verplanck, W. S., & Cotton, J. W. The dependence of frequencies of seeing on procedural variables: I. Direction and length of series of intensity-ordered stimuli. *J. gen. Psychol.*, in press.

(Received November 17, 1952)

Journal of Experimental Psychology
Vol 46, No. 1, 1953

THE PSYCHOPHYSIOLOGICAL SIGNIFICANCE OF THE GALVANIC SKIN RESPONSE [1]

A. C. MUNDY-CASTLE AND B. L. McKIEVER

National Institute for Personnel Research, South African Council for Scientific and Industrial Research

Since its discovery the galvanic skin response (GSR) has interested psychologists, but little clarity exists as to its significance. The historical background is covered by Prideaux (**32**), Farmer and Chambers (**10**), Woodworth (**40**), and McCleary (**28**).

The phenomenon is manifested in two different ways: (*a*) changes in skin resistance (Féré); and (*b*) differences in skin potential (Tarchanoff). There is a third category of similar origin, namely vasomotor changes, as measured by a plethysmograph.

These three are considered to be simultaneous and synchronous (**17, 18, 25, 36**). Jeffress (**25**) and Gildemeister (**17**) showed that similar latency periods and time relations exist in the Féré and Tarchanoff effects; Goadby and Goadby (**19**) were able to show that, although resistance and potential responses are different end-phenomena in the skin, they are activated and mediated by the same neural mechanisms since both are abolished by sympathectomy (**4, 19**). The resistance response, however, can be destroyed by total exsanguination of the limb, leaving the potential response unimpaired. They concluded that "the potential variations are expressions of activity at sympathetic nerve endings in blood vessels, sweat glands, etc. and changes in skin resistance and finger volume are due to vasoconstriction for which the sympathetic is the motor nervous pathway" (**19**, p. 182).

The GSR can be elicited by complex ideational stimuli, implying the involvement of high-level cortical structures, which are generally agreed to be in the premotor cortex (**12, 13**). Stimulation within this area causes complex manipulative movements in the distal portions of the extremities together with vasomotor and secretory changes (**7**).

With regard to subcortical mechanisms, most of the evidence points to a GSR center within the hypothalamus (see **28**).

The psychological significance of the GSR is controversial. Thouless (**35**) has suggested that resistance level reflects the preparedness of the organism to react; a British school headed by Aveling (**1**) holds that it is indicative of "conation." Woodworth (**40**) concludes that GSR is part of the emergency action of the sympathetic. According to Goadby and Goadby (**19**) the GSR is probably primitive in origin and is associated with preparation for activity and self-protection. The usefulness of the GSR as a preparatory and facilitative response for primitive man is stressed by Darrow (**6**).

Most often the GSR is considered as an indication of autonomic imbalance, albeit an inadequate one. Landis and Hunt (**27**) concluded that it was probably an indicator of change of direction of mental activity, but not a direct measure. That emotion is a concomitant of the GSR but is not wholly responsible for it is the conclusion at which Prideaux (**32**) finally

[1] The authors wish to thank Dr. S. Biesheuvel for advice and criticism.

arrived. The latency period of the response is considered by Golla (**20, 21**) as evidence in favor of the James-Lange theory of emotion. A positive correlation between rates of change in finger volume and emotional lability was found by van der Merwe and Theron (**34, 37**); this supports the claim that GSR is associated with emotion.

Biesheuvel (**2**), in his study on GSR, suggested that it is the somatic concomitant of Pavlov's Investigatory Reflex, which is also subject to habituation. The present experiment was undertaken at Dr. Biesheuvel's suggestion on the hypothesis that individual differences in GSR adaptation rate may be related to a temperamental factor underlying differences in excitatory/inhibitory balance.

Method

Measurement of skin resistance level (Féré effect) was by a Tinsley d.c. amplifier and ink pen recorder (**14**). The *S* lay on a soft couch in a moderately lit, partially soundproofed room; electrodes were attached to the palmar sides of the thumb and third finger of the left hand. Electrodes were of soft lead, shaped so that they could be moulded to the digits concerned, further security being ensured by elastic bands.

The skin beneath the electrodes was first cleaned with alcohol, then impregnated with electrode jelly. The *S* rested his left arm on his body and closed his eyes, these remaining shut during the whole experiment.

Auditory stimulation at 30-sec. intervals began after 2 to 3 min. Stimuli were produced by a notched disc rotating at a constant speed, thereby connecting a pair of earphones, via an isolating transformer to a 230 v. a.c. supply every 30 sec. Since the noise was harsh and loud, the earphones were placed on the head directed away from the ears in the reverse of the usual position. Each stimulus lasted for 1 sec. and presentation was continued until 3 successive stimuli had failed to elicit a GSR, or until 35 stimuli had been given and no adaptation had occurred. This number of stimuli was chosen for practical reasons, since it was difficult to retain *Ss* over an indefinite period. Both the controlled stimuli and any other stimuli likely to evoke GSR were marked on the record by *E*.

The *S* was warned beforehand that the noise would be loud and that the experiment might continue for as long as 30 min., although not necessarily so. A certain amount of continuous background hum arose from the motor driving the notched disc.

Possible variations due to polarization between the electrodes were checked by connecting them across a 10% saline solution and recording voltage changes over a period of 2 hr. No observable changes were recorded by the instrument.

The instrument used for recording skin resistance possessed a limitation in that recorder deflections corresponding to equivalent resistance changes decreased with increasing resistance. Consequently, variations of less than 500 ohms could not be accurately measured when the resistance level was 50,000 ohms or greater. This affected our results only slightly, since only five *Ss* showed such high resistance levels.

All *Ss* were given an electroencephalographic (EEG) recording by a six-channel Ediswan electroencephalograph, methods of recording and analysis being detailed in an earlier paper (**29**). For the purposes of this study, special note was made of frequency, amplitude, and per cent time of alpha rhythm (8 – 13 c/sec., parieto-occipital origin), and the incidence and amount of theta rhythm (4 – 7 c/sec.).

The experimental group comprised 109 *Ss*: 40 male and 33 female university students, 4 male and 26 female N.I.P.R. staff, 4 males and 2 females from the general population.

GSR records were assessed for the following: (*a*) the number of auditory stimuli before adaptation (if this occurred); (*b*) prestimulation resistance level; (*c*) the average resistance level, derived from four readings taken at equidistant quarters of the record during stimulation; (*d*) the extent of GSR to the first auditory stimulus (measured in ohms).

A separate pilot study of 36 normal *Ss* had revealed two distinct types of response, one occurring to external stimulation and the other occurring without stimulation; the latter were called endogenous responses, although some of them may have been due to unknown, extraneous stimuli. Similar spontaneous responses have been mentioned by other workers (**9, 27, 40**). This differentiation enables the classification of records as follows:

1. *Stable* (S Group): records showing no endogenous responses. Such records tended to maintain a constant trend in resistance level, i.e., a continued rise; an initial decrease followed by a gradual rise; or showing little variation. Responses were usually well defined and decreasing in amplitude as adaptation progressed (see Fig. 1).

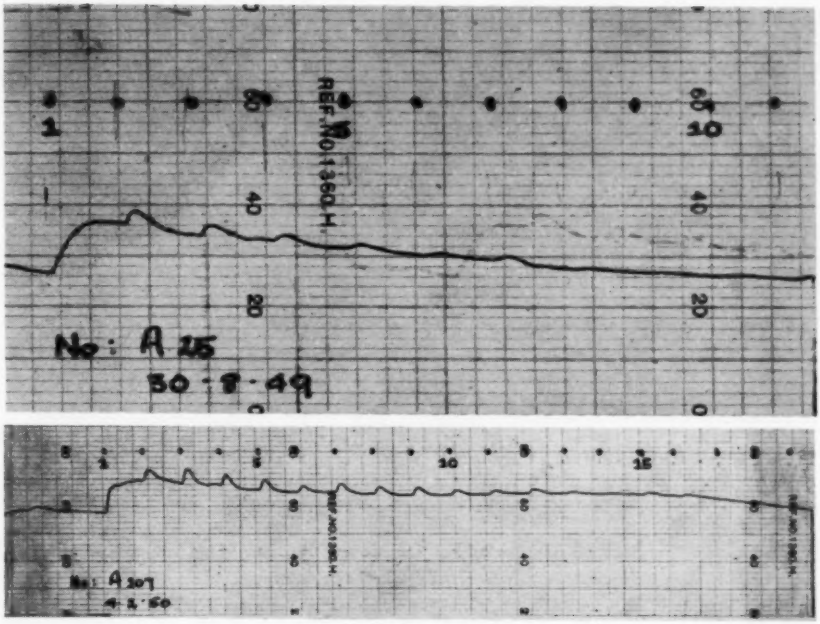

FIG. 1. Stable records

a. Male, 23 years. Adapted after 7 stimuli. Large first response and progressive decrease in response amplitude. Average resistance = 32,700 Ω.

b. Female, 22 years. Adapted after 16 stimuli. Similar to (*a*) but adaptation slower. Average resistance = 15,475 Ω.

2. *Labile* (L Group): records characterized by numerous endogenous responses, in addition to responses to stimuli. General resistance level often showed considerable variation while there was no consistent diminution in response amplitude as stimulation proceeded (see Fig. 2).

3. *Stable/Labile* (S/L Group). A third category is necessary in view of the occurrence of occasional dubious records showing isolated endogenous (not due to external stimuli) responses, or marked swings in resistance level (see Fig. 3).

This method of classification was adopted for the main experiment.

RESULTS

GSR adaptation rate.—Of the 109 *S*s, 8 were excluded from studies of adaptation rate for the following reasons: 5 *S*s showed average resistance levels above 50,000 ohms, making it impossible to assess type of GSR or adaptation, owing to the limitations of the machine already mentioned; 1 *S* showed no response at all to any stimulus; 1 *S* requested cessation of the experiment because of her feelings of acute distress; 1 *S* had to leave before the end of the experiment.

Composition of GSR groups and adaptation data are shown in Table 1.

From these figures it may be seen that all but two S records are accompanied by adaptation, whereas only six L records showed adaptation. It was therefore decided to determine the correlation between alpha frequency and adaptation rate, utilizing only S records, on the assumption that the L record reflected the intrusion of a different factor from the one under

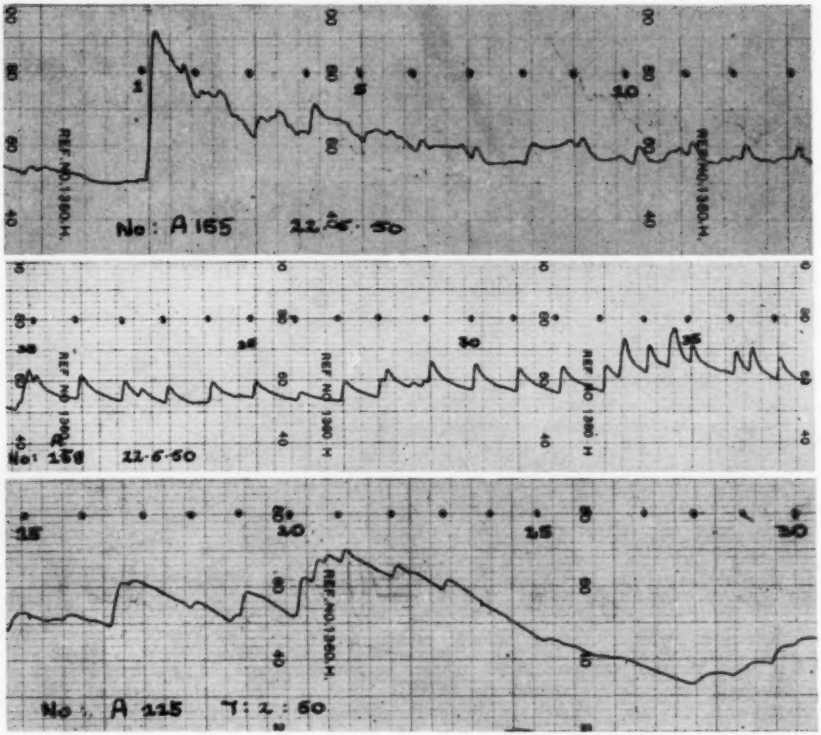

FIG. 2. Labile records

a. Female, 19 years. No adaptation. Very large initial response, numerous endogenous responses in addition to responses to stimuli. Generally irregular appearance. Average resistance = 30,352 Ω.

b. Female, 17 years. No adaptation. Responses show no decrease in amplitude and large endogenous responses occur. Average resistance = 11,025 Ω.

c. Female, 24 years. No adaptation. Marked swings in level of resistance and occasional endogenous responses. Average resistance = 23,850 Ω.

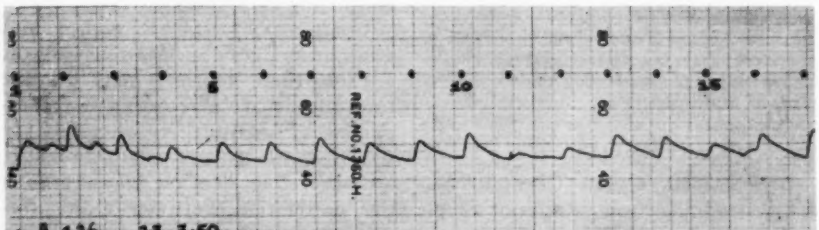

FIG. 3. Stable/Labile record

Male, 21 years. No adaptation. Small endogenous responses at first, but otherwise steady resistance level. Average resistance = 20,125 Ω.

TABLE 1

COMPOSITION OF GSR GROUPS

GSR Type	Total N	Population		Adaptation*		Sex	
		Staff	Nonstaff	Adapted	Did Not Adapt	M	F
Stable	47	17	30	45	2	23	24
Stable/Labile	21	4	17	11	9	9	12
Labile	35	6	29	6	28	14	21
Totals	103	27	76	62	39	46	57

*2 *Ss* excluded (see text).

consideration. It was not possible to include the two *S* records showing nonadaptation after 35 stimuli, since it cannot be known when or whether adaptation would have occurred.

The mean adaptation rate of the S group excluding these two *Ss* was 9.956 stimuli, $SD = 7.010$, and the product-moment correlation with alpha frequency is significant at the 1% level. Adaptation rate shows no relationship with age, average resistance level, or other alpha characteristics (Table 2).

Normality of the distributions of adaptation rate and mean alpha frequencies were checked by the Geary and Pearson (**16**) Test for Normality, with the following results:

Adaptation rate: $a = .727$

$\sqrt{b_1} = 1.654$

Alpha frequency: $a = .767$

$\sqrt{b_1} = .165$

TABLE 2

PRODUCT–MOMENT CORRELATIONS WITH ADAPTATION RATE
(S Group, $N = 45$)

	Mean	*SD*	*r*
Age	22.667	4.852	.022
Av. resistance level (ohms)	29,510.91	10,341.518	.073
α Frequency (c/sec.)	10.204	.860	.421
α Amplitude (microvolts)	33.167	21.766	−.136
% Time α	50.267	27.676	−.122

The high value of $\sqrt{b_1}$ in adaptation rate shows the distribution to be leptokurtic; none of the other values depart from normality. Since the distribution was not normal, the correlation of GSR adaptation rate of the S group and mean alpha frequency was also calculated by Kendall's rank-correlation method (**26**), giving the significant value:

$$\tau = .3979$$

$$\frac{(s-1)}{\sigma_s} = 3.503$$

$$N = 45.$$

Resistance level.—The prestimulation resistance level and the average resistance level were found to correlate significantly ($r = .791$, $p < .01$).

Since Darrow *et al.* (**8**) have reported relationships between conductance and alpha frequency and alpha amplitude, similar calculations with the present material were done, converting average resistance levels to micromho conductance values. There were no significant correlations. No relationships between per cent time alpha and average resistance levels were found (cf. **11**). In addition, there was no relationship between resistance level and age.

Types of GSR record.—Mean alpha frequencies of the three GSR types were not significantly different. Both mean age and mean average resistance level of the S group were significantly

higher than the L and S/L groups; the latter two groups not being significantly different from each other. There were significant differences in incidence of adaptation between all three GSR types. The figures are shown in Tables 3 and 4.

To check the possibility that lability was not being recognized at high resistance levels, owing to the limitations of the apparatus (see Method) and that the significant difference in mean resistance levels of the Labiles as opposed to the Stable group was not an artifact, all *S*s with resistance levels of 30,000 ohms or greater were excluded. The significance of the difference between the means of the two GSR types was recalculated. The resulting value of $t = 3.0$, $p < .01$ shows that the difference is still significant. Since in records with resistance values of 30,000 ohms and below, changes of the order of 200 ohms are easily recognizable, it is concluded that the differences found between L and S groups in the larger analysis are valid and not due to instrumental artifact.

χ^2 contingency tables showed no associations between GSR type and per cent time or amplitude of alpha rhythm, nor with sex. There was a significant association between GSR type and age (see Table 5). This finding, together with the fact that the nonstaff group was significantly younger than the staff group ($t = 2.5$ $p < .025$), prompted a test of the significance of the difference in incidence of lability in these two groups. It was possible that the nonstaff *S*s were more anxious and ill-at-ease in the test situation than the staff members, who were familiar with the surroundings and the *E*s (cf. **10**). However, the incidence of lability in the two groups is not significantly different ($t = 1.634$, $p > .10$).

GSR and theta rhythm.—Since GSR is associated with emotional variations,

TABLE 3

AVERAGE RESISTANCE LEVEL, ALPHA FREQUENCY, AGE, % INCIDENCE OF ADAPTATION IN GSR GROUPS

Group	Av. Resistance Level (ohms)			α Frequency (c/sec.)			Age			% Incidence of Adaptation
	Mean	*SD*	*N*	Mean	*SD*	*N*	Mean	*SD*	*N*	
S	29,072.5	10,497	47	10.437	.842	47	22.542	2.400	47	95.7
S/L	22,355.5	9,093	21	10.674	.930	21	19.928	1.419	21	55.0
L	21,986.5	9,768	35	10.617	1.065	35	20.786	1.988	35	17.6

TABLE 4

SIGNIFICANCE OF DIFFERENCES BETWEEN GSR GROUPS

Group	Mean Av. Resistance Level		Mean α Frequency		Mean Age		% Incidence of Adaptation	
	t	*p*	*t*	*p*	*t*	*p*	*t*	*p*
S & S/L	2.680	<.01	1.000	>.25	5.585	<.01	3.536	<.01
S & L	3.147	<.01	.826	>.25	3.613	<.01	10.121	<.01
L & S/L	.143	>.50	.210	>.50	1.877	>.05	2.899	<.01

TABLE 5

χ^2 CONTINGENCY TABLES: GSR TYPE AND ALPHA DATA, AGE, AND SEX

		S N	L N	S/L N	χ^2	p
% Time α Mean = 60.0	Above Mean	26	16	14	2.699	>.25
	Below Mean	20	18	6		
α Amplitude Mean = 35.69 microvolts	Above Mean	13	20	9	.290	>.75
	Below Mean	22	27	12		
Age Mean = 21.528	Above Mean	25	9	5	8.652	<.025
	Below Mean	22	26	16		
Sex	Male	14	23	9	.683	>.50
	Female	21	24	12		

the data were also analyzed in relation to theta rhythm in the EEG, as there is evidence that this is related to emotional expression (**29, 38**). No significant associations were found between incidence or amount of theta rhythm with average resistance level, size of GSR to the first auditory stimulus (assumed to be a partially emotional response), or GSR type.

DISCUSSION

Previous studies have indicated relationships between alpha frequency, cortical excitability characteristics, and certain temperamental aspects of behavior (**15, 30, 31**). Broadly speaking, high alpha frequencies occur with a high level of cortical excitability and tend to be accompanied by quick, impulsive, and variable behavior; low alpha frequencies occur with low levels of cortical excitability and are accompanied by slow, cautious, steady behavior. These aspects of behavior are comparable with the two poles of the Heymans-Wiersma temperament variable, Secondary Function (**3, 22, 39**).

The significant correlation between alpha frequency and the rate of adaptation of the Stable group suggests that adaptation rate in this group may also be related to cortical excitability level. Further research must be undertaken before it can be ascertained whether there is a definite relationship with the temperament variable Secondary Function. There is no correlation with alpha frequency when the Labile group is included, nor is there any significant difference in mean alpha frequency between the three GSR classifications, this indicating the operation of more than one factor.

The habituation or extinction of GSR in S records may be explained by the development of internal inhibition, as postulated by Pavlov in his work on conditioning, particularly since disinhibition may occur if a novel stimulus is introduced. It was noticed that spontaneous recovery of previously extinct responses occurred after a period of time. For example, a person who shows adaptation after 15 stimuli may respond after a break of 24 hr. to 10 stimuli before readaptation. This means that the relationship between alpha frequency and adaptation rate in S records will only be observed if experimental conditions are new to *S*. It would therefore appear that one of the factors in these results involves the relative strength of excitatory/inhibitory processes. If the weight is on the excitatory side, adaptation is slow; if on the inhibitory side, it is quick, other conditions being equal. This conclusion puts GSR adaptation rate in S records in the same context as alpha frequency and Secondary Function.

It is now necessary to consider why the relationship does not hold when L and S/L groups are included. The most significant findings are: (*a*) A significant association between GSR and age; young people showing predominantly L and S/L records. (*b*) The average resistance levels of the L and S/L groups are significantly lower than that of the S group. There was no association between resistance level and age. (*c*) In-

cidence of adaptation was significantly greater in the Stable than in the L and S/L groups.

The association between GSR groups and age immediately suggests that stability may be a function of maturity; there is increasing evidence to show that there are centers for autonomic control in the frontal cortex. A wide variety of autonomic reactions has been elicited by electrical stimulation of the frontal lobes, among which have been GSR in the cat (**28, 33**) and vasomotor reaction in man (**13**). Fulton (**12, 13**) regards the recognition of the relationships between the forebrain, autonomic function, and affective behavior as one of the most significant developments in modern neurology. These findings support the suggestion that lability of GSR records might be due to relative lack of cortical control over lower centers involved in this sphere of autonomic function, this possibly reflecting a degree of cortical immaturity. Such a conclusion accords with Darrow's theory (**7**) that the frontal cortex exerts inhibitory restraint over hypothalamic activity, this restraint being removed by premotor injury. It should be noted, however, that there was no observed relationship with theta rhythm, which has been shown to be related to cortical maturation processes (**23, 29**). This does not invalidate the above suggestions since EEG's and GSR's were not recorded simultaneously and there was no automatic frequency analysis. Consequently, only gross assessments of theta activity were possible.

Variations in extent of cortical control imply corresponding variations in thresholds for the elicitation of changes in skin resistance; this means that *Ss* with L records will show GSR's to stimuli which are ineffective with *Ss* with S records, other conditions being equal. Therefore it would be expected that general levels of resistance will be lower in the L than S groups, and, further, that any process of adaptation to a repetitive stimulus may be interfered with by the heightened reactivity and its consequent disinhibiting effects. Both expectations are confirmed.

There seem to be, therefore, two factors underlying the results of this experiment. The first factor determines the relative strength of excitatory/inhibitory processes, which is reflected in adaptation rate in Stable records. It has been shown to be related to alpha frequency, hence to central excitability characteristics and possibly the temperament variable Secondary Function (**15, 30, 31**). The second factor, manifested in lability, appears to be related to the degree of cortical restraint over lower autonomic centers. In *Ss* with a high degree of lability the relative excitatory/inhibitory balance is not reflected in adaptation rate, since the adaptation process is continually disinhibited by the increased responsiveness consequent on lack of cortical restraint. The factor giving rise to this disinhibitory process is unrelated to the excitatory/inhibitory factor, but its presence may indicate apparent high position on the excitation side because of failure in adaptation.

It is impossible to ascertain the distribution of lability from the data. We postulate, however, that lability is a unitary continuous variable, of which progressively less is present in the S/L and S groups. The fact that there was no significant difference in incidence of lability between staff and nonstaff groups despite the lower mean age of the latter, shows that lability is not necessarily a function of unfamiliarity. It may be related to another temperament variable such as emotionality or emotional stability.

Although no significant relationships were found between GSR and other EEG data, this does not mean that they do not exist, nor does it invalidate the findings of Darrow (**8**) mentioned earlier. Such relationships can probably only be found by simultaneous recording of GSR and EEG. In particular, the relationships between theta rhythm in EEG and GSR lability should be of interest since the former has been shown to be related to cortical immaturity (**23, 24**). An experiment involving simultaneous EEG and GSR recording has been planned.

Summary and Conclusions

1. GSR's of 109 normal *Ss* were recorded while auditory stimuli were given at 30-sec. intervals. Recordings continued until adaptation occurred or until 35 stimuli had been presented. EEG's were independently recorded from all *Ss*.

2. GSR records were classified as follows:

a. Stable (S): GSR's only to observed stimuli: no endogenous responses. Consistent pattern of resistant changes.

b. Stable/Labile (S/L): a few endogenous responses and/or unstable resistance level.

c. Labile (L): many endogenous responses, often accompanied by unstable resistance level.

3. Labiles tended not to adapt to the stimuli, whereas Stables did, their adaptation rate being significantly correlated with alpha frequency. Mean

age of L group was significantly lower than that of S group.

4. It was suggested that two factors contributed to these results:

a. A factor determining the relative strength of excitatory/inhibitory processes, this also related to alpha frequency and probably the temperament variable Secondary Function.

b. Extent of cortical restraint over lower autonomic centers, this probably related to a maturation factor. Adaptation rate will be determined by the first factor, but if the second factor is predominant, an L record results and continued disinhibition of the adaptation process occurs.

5. A further experiment recording GSR and EEG simultaneously has been planned to investigate thoroughly these hypotheses.

References

1. Aveling, F. The psychogalvanic phenomenon. *Proc. IX int. Congr. Psychol.*, 1929, **9,** 63–65.
2. Biesheuvel, S. Unpublished master's thesis, Univer. of Cape Town, 1930.
3. Biesheuvel, S. An observational technique of temperament and personality assessment. *Bull. nat. Inst. Pers. Res.*, 1949, **1,** No. 4, 9–27.
4. Carmichael, E. A., Honeyman, W. M., Kolb, L. C., & Stewart, W. K. A physiological study of the skin resistance response in man. *J. Physiol.*, 1941, **99,** 329–337.
5. Darrow, C. W. Sensory, secretory and electrical changes in the skin following bodily excitation. *J. exp. Psychol.*, 1927, **10,** 197–226.
6. Darrow, C. W. The palmar galvanic skin reflex (sweating) and parasympathetic activity. *Amer. J. Psychol.*, 1936, **48,** 522–524.
7. Darrow, C. W. Neural mechanisms controlling the palmar galvanic skin reflex and palmar sweating. *Arch. Neurol. Psychiat.*, 1937, **37,** 641–663.
8. Darrow, C. W., Pathman, J., & Kronenberg, G. Level of autonomic activity and electroencephalogram. *J. exp. Psychol.*, 1946, **36,** 355–363.
9. Deutsch, F. Thus speaks the body. II. *Acta Medica Orientalia*, 1950, **IX,** 199–215.
10. Farmer, E., & Chambers, E. G. Concerning the uses of the psychogalvanic reflex in psychological experiments. *Brit. J. Psychol.*, 1924–1925, **15,** 237–254.
11. Freeman, G. L. Cortical autonomous rhythms and the excitation levels of other bodily tissues. *J. exp. Psychol.*, 1940, **27,** 160–171.
12. Fulton, J. F. *Functional localization in the frontal lobes and cerebellum.* London: Oxford Univer. Press, 1949.
13. Fulton, J. F. *Frontal lobotomy and affective behavior.* New York: W. W. Norton, 1951.
14. Gall, D. C. A direct-current amplifier and its application to industrial measurements and control. *Brit. J. Inst. elect. Eng.*, 1942, **89,** 434–443.
15. Gastaut, H. et Y., Roger, A., Corriol, J., & Naquet, R. Étude électrographique du cycle d'excitabilité cortical. *EEG clin. Neurophysiol.*, 1951, **3,** 401–428.
16. Geary, R. C., & Pearson, E. S. *Tests of normality.* Biometrika Office. Cambridge Univer. Press, 1938.
17. Gildemeister, M. *Handbk. normal path. Physiol.*, 1928, **8,** 657–702, 766–784.
18. Gilliat, R. W. Vasoconstriction in the finger after deep inspiration. *J. Physiol.*, 1948, **107,** 76–88.
19. Goadby, K. W., & Goadby, H. K. The nervous pathway of the psychogalvanic response. *J. Physiol.*, 1949, **109,** 177–184.
20. Golla, F. Croonian lectures. *Lancet*, 1921.
21. Golla, F. In N. G. Harris (Ed.), *Modern trends in psychological medicine.* London: Butterworth Medical Publication, 1948. Pp. 105–121.
22. Heymans, G. *Inleiding Tot de speciale Psychologie.* Haarlem: De Erven. F. Bohn, 1929. 2 vols.
23. Hill, D. Cerebral dysrhythmia; its significance in aggressive behavior. *Proc. roy. Soc. Med.*, 1944, **37,** 317–330.
24. Hill, D., & Watterson, D. Electroencephalographic study of psychopathic personalities. *J. Neurol. Psychiat.*, 1942, **5,** 47–65.
25. Jeffress, L. A. Galvanic phenomena of the skin. *J. exp. Psychol.*, 1928, **2,** 130–144.
26. Kendall, M. G. *Rank correlation methods.* London: Charles Griffin, 1948.
27. Landis, C., & Hunt, W. A. The conscious correlates of the galvanic skin response. *J. exp. Psychol.*, 1935, **18,** 505–529.

28. McCLEARY, R. A. The nature of the galvanic skin response. *Psychol. Bull.*, 1950, **47**, 109–124.

29. MUNDY-CASTLE, A. C. Theta and beta rhythm in the electroencephalograms of normal adults. *EEG clin. Neurophysiol.*, 1951, **3**, 477–486.

30. MUNDY-CASTLE, A. C. An analysis of central responses to photic stimulation in normal adults. *EEG clin. Neurophysiol*, 1953, **5**, 1–22.

31. MUNDY-CASTLE, A. C. EEG in relation to temperament. *4th Annual Congress of the S. A. Psychol. Association.* Pietermaritzburg, 4th July, 1951. Submitted for publication.

32. PRIDEAUX, E. The psychogalvanic reflex: a review. *Brain*, 1920, **43**, 50–71.

33. SCHWARTZ, H. C. Effect of experimental lesions of the cortex on the psychogalvanic reflex in the cat. *Arch. Neurol. Psychiat.*, 1937, **38**, 308–320.

34. THERON, P. A. Peripheral vasomotor reactions as indices of basic emotional tension and lability. *Psychosom. Med.*, 1948, **10**, 335–346.

35. THOULESS, R. H. The causes of the continuous changes of resistance observed in psychogalvanic experiments. *Brit. J. Psychol.*, 1925, **16**, 5–15.

36. UHLENBRUCK, P. *Z. Biol.*, 1924, **81**, 51–56. Referred to by Woodworth (**40**).

37. VAN DER MERWE, A. B., & THERON, P. A. A new method of measuring emotional stability. *J. gen. Psychol.*, 1947, **37**, 109–124.

38. WALTER, W. G. The twenty-fourth Maudsley lecture: The functions of electrical rhythms in the brain. *J. ment. Sci.*, 1950, **96**, 1–31.

39. WIERSMA, E. D. *Lectures on psychiatry.* London: H. K. Lewis, 1932.

40. WOODWORTH, R. S. *Experimental psychology.* New York: Henry Holt, 1938.

(Received November 3, 1952)

Journal of Experimental Psychology
Vol 46, No. 1, 1953

PHYSIOLOGICAL NEED, WORD FREQUENCY, AND VISUAL DURATION THRESHOLDS [1]

LAUREN G. WISPÉ AND NICHOLAS C. DRAMBAREAN

The Ohio State University

During the past 15 years a growing body of research has demonstrated a significant relationship between various physiological need states and certain aspects of perception. Despite differences in methods, various studies (**2, 4, 10, 15, 16**) have shown that in so-called "projective" experimental situations *Ss* report more need-related activities and objects as a function of increased deprivation states. More recently several experiments (**13, 14**) have used the tachistoscopic exposure of words related to the six value areas of the Allport-Vernon Scale of Values (**1**) as a means of studying the differential effects of personal values in perception. Tachistoscopic-recognition time studies (**13, 14**) have in common with the projection-type studies (**2, 10**) the presentation of only minimal cues to *Ss*. Although this similarity applies only to the incipient stages of the tachistoscopic procedures, the results of these two methods are roughly comparable, and interesting possibilities present themselves for an amalgamation of the two techniques. It is still questionable that a temporary physiological need state will selectively affect the tachistocopic recognition of need-related word stimuli. The data from the projection experiments indicate that need states are important determiners of what *S* will report in the experimental situation, and the results of the tachistoscopic studies show that values will selectively influence the recognition of words related to these value areas. Therefore, in the light of previous experimental evidence, it was hypothesized that with an increasing physiological need state the recognition-time thresholds for need-relevant verbal stimuli would be lowered, or that hungry and thirsty *Ss* would "see" food and water words more rapidly than neutral words.

Since changes in perception are inferred from changes in verbal response, the properties of linguistic materials must be taken into account in any analysis of perceptual phenomenon. The strength, or "frequency" in Thorndikian terms, of the word stimuli may well be a factor that is contributing to what is inferred in the perceptual process. Howes and Solomon (**8**) investigated this relationship and found the frequency of usage of words to be inversely related to the duration thresholds for the recognition of words. These authors assumed that duration thresholds and verbal frequency are both related to verbal strength, or the probability that the verbal response will occur. The present study was undertaken to confirm the Howes-Solomon hypothesis that common words would be recognized more rapidly than uncommon words.

The two points of view presented above, one emphasizing autistic factors in perception and the other indicating the importance of the stimulus variables, lead to certain important problems. Murphy has stated (**11,**

[1] The Laboratory of Social Relations, Harvard University, supported the pilot study. The present research was made possible by a grant from The Ohio State University Research Foundation. The writers are greatly indebted to Dr. R. L. Solomon for his assistance in the preliminary stages of this project. We would also like to thank Drs. R. S. Wherry, D. D. Wickens, and P. Ratoosh for many helpful suggestions and criticisms.

p. 367) that in the face of relatively inflexible uniformities of structure (in this case the frequency of word usage in our language) autistic factors may often be negligible. And Solomon and Howes (**17**), repeating the Postman, Bruner, and McGinnies (**13**) experiment using common and uncommon words in the Allport-Vernon areas, found that word-commonness was a more important determinant of recognition threshold than personal values. In fact, Solomon and Howes prefer to consider personal values as mean deviations from the normal-population verbal usage. Postman and Schneider (**14**), on the other hand, have argued against S–R reductionism and for the value of including "directive" factors among the theoretical considerations. The crucial question, therefore, is this: How important is a temporarily established need state in the face of strongly conditioned stimulus-response relations? Will the verbal responses to a particular word stimulus, responses which have been conditioned through the years, be significantly altered by an induced physiological need? Or, to put it specifically, after 24 hr. without food and water, will the name of an uncommon thirst satisfier be recognized more rapidly than a common neutral word?

Method

Stimulus material.—The stimulus material consisted of two lists of 24 words each, 12 of which were "need" words and 12 of which were "neutral" words. Of the 12 need words in each list, 6 related to food deprivation and 6 to water deprivation. Following the lead of McClelland and Atkinson (**10**), the food and water need words were divided into "act" words, which in the present study signified the process of taking food and/or water into the body, and "object" words, which were the names of satisfiers. There were thus four subclasses of need words: (*a*) food-act words, (*b*) food-object words, (*c*) water-act words, and (*d*) water-object words. The 12 neutral words in each list were matched for commonness and supposedly were unrelated to the hunger-thirst process. Both lists were matched for commonness, and, insofar as possible, contained the same number of act and object, food, water, and neutral words. Testing time for each list was about 1½ hr.

Several independent measures of word frequency were used in order to assess the commonness of the word stimuli. The Thorndike-Lorge *Teacher's Word Book of 30,000 Words* (**18**), which contains indications of word strength based upon frequency of occurrence in their sample, was the basic measure. However, the Thorndike-Lorge estimates are based on a selected sample of written language, and it was believed that there might be some disparity between this and the spoken language of college undergraduates. The problem of ascertaining commonness is aggravated by the fact that in the construction of the word lists little latitude is given by the structure of our language; to wit, there are just a limited number of words signifying the process of taking food and water into the body. With these facts in mind the stimulus words were presented to two samples of 150 and 100 college undergraduates to be rated as "common," "mid-common," or "uncommon" depending upon the frequency with which they are used in the everyday vocabulary of the judges. The correlations between the Thorndike-Lorge word ratings and the commonness ratings of the undergraduate samples were .86 and .81, respectively. In view of these high correlations with the Thorndike-Lorge ratings, the verbal frequency estimates of the students were used to determine the commonness of the words used in the experiment.[2] Table 1 included the ratings and classifications of the words used.

Equipment.—Presentation of the stimulus material was made with an episcotister and a 50-mm. slide projector. The episcotister was driven at the rate of one revolution in 3.5 sec. by a constant-speed motor. The disk was 18¼ in. in diameter, and the arc openings were calibrated so that exposures could be graduated at .01-sec. intervals. The words were typed in regular pica type on amber-colored cellophane and mounted on 2 × 2-in. slides. The *S*s sat 15 ft. from the beaded glass screen on which two fixation points had been painted. Five feet behind the *S*s, and slightly to their right, was the projection table.

Subjects.—The 60 *S*s, Ohio State University men and women enrolled in the introductory psychology course, all volunteered thinking they would have to go 24 hr. without food or water. Since the pilot study indicated that 24-hr. abstinence produced mild fatigue effects which seemed to be aggravated by having *S* reappear for several

[2] As a further check on the reliability of the ratings for commonness of usage the experimental *S*s rated the words for frequency of usage about two weeks after they had completed the recognition study. The correlations with the Thorndike-Lorge and the two undergraduate ratings were .79, .89, and .97, respectively.

TABLE 1

PRESENTATION ORDER AND CLASSIFICATIONS OF LISTS OF STIMULUS WORDS WITH MEANS AND *SD*'S

List I				List II			
Stimulus Word	Classification*	Mean	*SD*	Stimulus Word	Classification*	Mean	*SD*
munch	AUF	.360	.065	nibble	AUF	.242	.152
godson	NU	.312	.312	devout	NU	.171	.093
need	NC	.147	.055	beat	NC	.106	.036
mean	NC	.317	.173	speak	NC	.118	.052
serenade	NU	.167	.071	solar	NU	.186	.132
feed	ACF	.107	.058	eat	ACF	.158	.100
lemonade	OUW	.129	.064	soda	OUW	.162	.076
think	NC	.074	.026	rip	NC	.124	.060
comb	NC	.126	.057	weather	NC	.194	.086
imbibe	AUW	.280	.163	gulp	AUW	.123	.073
milk	OCW	.112	.059	cake	OCF	.088	.043
imbed	NU	.203	.116	gust	NU	.117	.080
drink	ACW	.061	.021	sip	ACW	.108	.062
coke	OCW	.096	.071	water	OCW	.113	.062
dine	ACF	.103	.061	drain	ACW	.086	.033
waddle	NU	.234	.104	idler	NU	.187	.117
waffle	OUF	.147	.139	cider	OUW	.107	.041
chuckle	NU	.182	.093	ram	NU	.184	.062
mine	NC	.106	.042	lake	NC	.079	.026
hunch	NU	.202	.109	quibble	NU	.129	.051
chocolate	OUF	.164	.106	ham	OUF	.130	.053
guzzle	AUW	.130	.089	devour	AUF	.079	.025
dive	NC	.079	.043	dread	NC	.118	.064
meat	OCF	.179	.093	steak	OFC	.098	.038

* A, O, N indicate "act," "object," and "neutral" words, respectively; C, U indicate "common" and "uncommon" words; and F, W indicate "food" and "water" words.

protracted testing sessions, different *S*s were used for each of the three deprivation periods, and each *S* was tested only once. The *S*s were randomly assigned to one of the two lists and one of the three deprivation periods. Thus each list at each of the deprivation periods was run on 10 *S*s.

Procedures.—The experiment was presented as one of visual acuity and deprivation. On the testing days *S*s were not permitted to eat or drink, and they were requested to limit their activities to classes and study.[3] Smoking, if necessary, was permitted with restraint. Three periods of deprivation were induced. The control group (0–2-hr. deprivation) ate a normal, but late, lunch. The 10-hr. deprivation group ate breakfast at 7:30 A.M., and the 24-hr. deprivation group ate their last meal at dinner time on the night preceding their testing periods. All testing was begun each day at 5:30 P.M. Each *S* was called the night before he was to appear in order to insure his abiding by the conditions.

In the instructions use of the episcotister was explained to S, and he was told that the list of words to be presented had neither proper names nor nonsense syllables. The *E* explained that "Ready?" would precede each presentation, and that when nothing was clear, *S* should report "flash" or "blank." Guessing was encouraged.

The two practice words were intended not

[3] The question of a food, or a water, or a combined deprivation state provides additional complications for research of this kind. The food drive in our culture is cyclical, and largely a matter of habit. Other research (16) has indicated that the factor of habit and the influence of a real physiological need state are almost inseparable. The water drive, although relatively free of the above-mentioned problems, is difficult to arouse. Probably 24 hr. of water deprivation induces little real tissue need in the average person. Furthermore, in experiment (9, 19) and experience of most people the two drives can be shown to be almost inextricably bound together. Nearly everyone has sometime taken certain of the water-objects to satisfy the hunger need—as when he drinks a great deal of water to ward off hunger pangs. There is the final problem that it is almost impossible to arouse one need state without arousing the other also. In order to create a water need it is necessary to deprive the person not only of water, but also of certain foods which are high in water content. These considerations suggest that the best procedure was to deprive *S*s of both food and water.

TABLE 2

MEAN RECOGNITION TIME THRESHOLDS AND *SD* (IN SEC.) OF PRINCIPAL CLASSIFICATIONS OF STIMULUS WORDS

Deprivation Periods		Need Words		Neutral Words	
		Common	Uncommon	Common	Uncommon
0	*M*	.151	.241	.129	.189
	SD	.048	.080	.041	.059
10	*M*	.089	.136	.128	.188
	SD	.028	.045	.045	.058
24	*M*	.089	.138	.142	.192
	SD	.064	.048	.047	.062

only to absorb some of the practice effects, but also to give *E* an opportunity to estimate the range within which *S*'s first correct responses would occur. Since this range was extremely large, no definite sequence of exposure times could be used. Therefore, for each *S* the exposure time of the first set of presentations was aimed at the level where *S* would report seeing "only a flash of light." The exposures were then increased at .03-sec. intervals until *S* could report seeing "letters" and/or began making some guesses about the word. Thereafter the exposure times were increased at .01-sec. intervals until the word was correctly reported. In this ascending method of limits, two presentations were made at each exposure level. The criterion for a correct report was two successively correct responses. The duties of *E* forced the intervals between presentations to remain fairly constant at about 25 sec. During this time *S*s were allowed to look around the room, which was kept at about 5–7 ft.-candles illumination.

RESULTS

Preliminary to the analyses of variance, the mean recognition time for List I was tested against the mean recognition time for List II. List I turned out to be significantly more difficult, but the shape of the various functions within each list was the same, so the data from the two lists were combined. To test the assumption of homogeneity necessary for the analysis of variance, the Bartlett test was applied to the combined data. The result, $\chi^2 = .1202$, $p > .05$, indicates that these data fulfill this assumption.

Since several analyses of variance failed to reveal any significant differences between the food-water and the act-object mean recognition times, the data were classified into (*a*) need-neutral words, (*b*) common-uncommon words, and (*c*) the three periods of deprivation. These mean recognition times are given in Table 2. From Fig. 1 we see that in general the need words are recognized more rapidly as need increases, while the recognition times for the neutral words are unrelated to the periods of deprivation. The analysis of variance of these data (Table 3) shows that of the main effects only the differences between the common- and uncommon-word recognition thresholds proves significant. When properly tested neither the difference between the need and neutral words, nor that among the deprivation periods is significant.

The only significant interaction effect, and the one most important for our purposes, is between the need and neutral words and the three periods of deprivation. This interaction is significant beyond the 1% level. The *t* tests for the significance of differences between the mean recognition times of

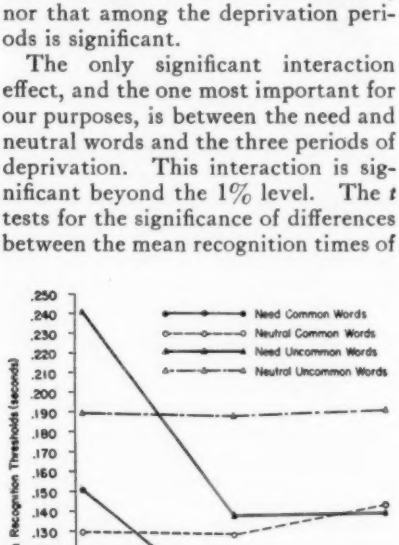

FIG. 1. Recognition thresholds (in seconds) for the need and neutral words of the three deprivation periods

TABLE 3

ANALYSIS OF VARIANCE OF RECOGNITION TIME THRESHOLDS

Source	*df*	Variance Estimate	*F*
Word commonness	1	12.6914	27.5*
Deprivation periods	2	2.53819	
Need-Neutral words	1	1.55418	
Commonness × Deprivation	2	.23049	
Commonness × Need-neutral words	1	.02176	
Deprivation × Need-neutral words	2	3.00214	6.5*
Commonness × Deprivation × Need-neutral words	2	.14445	
Within cells	132	.46093	
Total	143		
W. C. + C × D × N + C × N + C × D	137	.4497	

* $p < .01$.

the need and neutral words and the deprivation periods show that the results are in the expected direction; at 0-hr. there are no significant differences between the need and the neutral words, while at the 10-hr. and 24-hr. deprivation periods the need-word thresholds are significantly lower than the neutral-word thresholds. The difference between the 10-hr. and the 24-hr. need-word recognition times, however, is not significant.

It is of interest to separate the factors causing variability in recognition times, and to arrive at estimates of the components of the total variance assignable to each factor.[4] Essentially, we are asking the question, which are the more important components of variance? Computing these, and their standard errors, we see that the largest part of the variance is attributable to commonness, $\hat{\sigma}_C = .1703 \pm .0215$. However, the interaction between the deprivation periods and the need-neutral words is also large, $\hat{\sigma}_{D\times N} = .1062 \pm .0079$. The components for the deprivation periods, and for the need-neutral words, are $\hat{\sigma}_D = .0096 \pm .0035$, and $\hat{\sigma}_N = .0202 \pm .0012$. Although commonness is the most important factor, the estimate of the variance ascribable thereto is only about 1½ times greater than that contributed by the deprivation × need-word interaction.

The *S*s' presolution responses were recorded and analyzed. In general these data confirm the hypothesis (**13**) of an inverse relationship between the number of "covaluant" presolution responses and the recognition-time thresholds, for there is an increase in the number of need-related presolution responses at 10 hr. with a corresponding decrease in the duration thresholds, while at 24 hr. the number of presolution responses decreases and the recognition times increase.

Important information can be obtained if the presolution responses are categorized into those denoting "goal objects" (the names of hunger and thirst satisfiers and the names of meals), "goal acts" (verbs and verb forms related to taking foods and liquids into the body), "instrumental responses" (names of specific establishments serving food and beverages, names of cooking and containing utensils, and names of processes for the preparation of food) and "neutral responses" (all those not classifiable under one of the rubrics above). These results are presented in Fig. 2. There is a sharp increase in the number of act and object responses at 10 hr., and a marked drop at 24 hr. The instrumental responses, on the other hand, drop slightly at 10 hr., and rise slightly at 24 hr. The χ^2 analysis of the data is significant beyond the 1%

[4] The experimental design employed in this research is better suited to specifying the fixed relations of the parameters of interest than to estimating the components of random variation, especially since the procedure employed deviated somewhat from a strictly random sampling of the variables involved (**5, 7**). Nevertheless, as Daniels (**6**) has pointed out, one may interpret the mean squares of fixed deviations as variance components, and proceed to assign portions of the total variance to each of the factors involved.

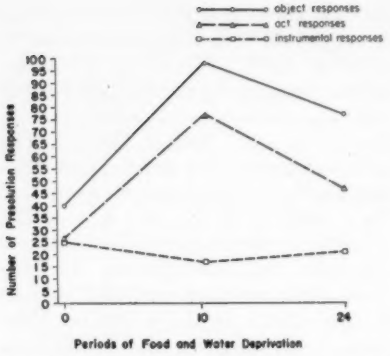

FIG. 2. Number of presolution responses for each deprivation period

level. This empirical finding is corroborated by the results of word-association data (**20**), and it is related to McClelland and Atkinson's findings (**10**) that as need increases there is an increase in the number of responses pertaining to acts instrumental to obtaining need-satisfiers.

DISCUSSION

Let us first consider the findings on physiological need states and recognition times for need-neutral stimuli. In general we can conclude that with increasing need, need-related words are recognized more rapidly. However, this is not a simple linear relationship, for by the 24-hr. period there is a slight (although not significant) increase in the recognition thresholds. Since several analyses of variance indicate this leveling-off effect is not the result of "subjective" hunger and thirst states (measured by two 5-point rating scales), and since it is not due to different food- and water-word effects, we can assume the shape of this curve to be a function of the processes under investigation.

Various theoretical explanations can be made of these findings. In learning theory terms, we can say that since in the past history of the organism the food and water objects and acts named in the word list have been associated "subvocally" with the reduction of the food and water drives, their probability for occurrence is raised in the response hierarchy. Since associationist learning theories hold that the more frequently a response and a stimulus have been associated in the past, the briefer the exposure necessary to elicit that response, it follows that the overt responses, the recognition of food and water words, will be elicited by lower presentation thresholds. This is in keeping with the explanation of Solomon and Howes, who write that "the relation between word frequency and duration threshold is a case of the general relationship between the probability of a response and the latency of its occurrence in the presence of discriminated stimulus" (**17,** p. 267).

On the other hand, in terms of the Bruner-Postman "hypothesis-information" theory (**3, 12**) the results can be explained by the proposition that "the stronger the motivational support for a given hypothesis, the less of appropriate stimulus information is required to confirm it" (**12,** p. 255). In the present experiment "motivational support" could be assumed to have increased with increasing deprivation, and "less stimulus information" can be equated to shorter presentation exposures.

This raises the question of the effect of word frequency on recognition thresholds. The problem has not always been clearly put. For example, Solomon and Howes have shown that a person's score on the Allport-Vernon Scale of Values can be deduced from "the inverse relation between word frequency and duration threshold, plus a knowledge of how frequently that subject uses words representing the various areas of interest" (**17,** p. 266). These writers presented a strong argument showing that more of the variance in recognition thresholds could be accounted for by word-commonness than by such variables as *S*'s value rank. But this is not the issue. In the present research, too, the largest component of variance is attributable to commonness. However, even in the Solomon and Howes study there is a marked relationship between value rank and duration threshold, especially for uncommon words; and in the present experiment the interaction between the need words and deprivation is highly significant. Therefore it would appear that although word-commonness may be the dominant source of variance, much more variance in this kind of research can be explained by including so-called "motivational" or "directive" variables.

One further question remains, and this concerns the role of word frequency in the interaction between the need-neutral words and the deprivation periods. The question is: Will common and uncommon need and neutral words behave in the same way under increasing periods of deprivation? Studies already cited (**14, 17**) have indicated the importance of this aspect of the problem. An adequate theory of perception must consider not only the stimulus and the "directive" variables, but must also account for the dynamics of the interaction between them. Figure 1 shows that both the common and the

uncommon need words are recognized more rapidly with increasing deprivation, although the curve for the uncommon words seems to have a steeper slope between the 0- and the 10-hr. deprivation periods. The statistical significance of this relationship is tested by the triple interaction in Table 3, which is not significant. Thus we may conclude that in this research word-commonness was not an important factor in the relationship between the need-neutral words and the periods of deprivation.

Summary

This experiment was designed to test the hypothesis that with increased deprivation, need-relevant stimulus words would be recognized more rapidly than neutral words, and to determine what part word frequency played in this process. Three groups of 20 *S*s each were deprived of food and water, and their recognition thresholds for need-relevant and neutral words were obtained at 0-, 10-, and 24-hr. deprivation. The stimulus words were matched for commonness and food-water relevance. The principal findings were: (*a*) that word-commonness was the single most important factor, (*b*) that need-related words were recognized more rapidly as need increased, and (*c*) that word-commonness was not a significant interactive factor. The relationship between physiological deprivation and recognition-time thresholds for need-relevant stimuli is not a simple linear one, however, for there is a slight increment in the recognition-time thresholds at the 24-hr. period. Analysis of the presolution responses revealed that at 10 hr. there was an increase in the number of need satisfiers and "eating" and "drinking" words, with a drop at 24 hr.; while the responses denoting acts instrumental to obtaining need-satisfaction decreased at 10 hr. and increased at 24 hr.

References

1. Allport, G. W., & Vernon, P. E. *A study of values*. Boston: Houghton-Mifflin, 1931.
2. Atkinson, J. W., & McClelland, D. C. The projective expression of needs: II. The effect of differential intensities of the hunger drive on thematic apperception. *J. exp. Psychol.*, 1948, **38**, 643–655.
3. Bruner, J. S. Personality dynamics and the process of perceiving. In R. R. Blake & G. V. Ramsey (Eds.), *Perception, an approach to personality*. New York: Ronald, 1951.
4. Chein, I., Levine, R., & Murphy, G. The relations of the intensity of a need to the amount of perceptual distortion. *J. Psychol.*, 1942, **13**, 283–293.
5. Crump, S. L. The estimation of variance components in analysis of variance. *Biometrics Bulletin*, 1946, **2**, 7–11.
6. Daniels, H. E. The estimation of components of variance. *Suppl. J. roy. Statist. Soc.*, 1939, **6**, 186–197.
7. Eisenhart, C. The assumptions underlying the analysis of variance. *Biometrics*, 1947, **3**, 1–21.
8. Howes, D. H., & Solomon, R. L. Visual duration thresholds as a function of word probability. *J. exp. Psychol.*, 1951, **41**, 401–410.
9. Lindsley, O., & Verplanck, W. S. Water ingestion as a function of previous food and water deprivation. *Amer. Psychologist*, 1952, **7**, 257. (Abstract)
10. McClelland, D., & Atkinson, J. The projective expression of needs: I. The effect of different intensities of the hunger drive on perception. *J. Psychol.*, 1948, **25**, 205–222.
11. Murphy, G. *Personality*. New York: Harper, 1947.
12. Postman, L. Toward a general theory of cognition. In J. H. Rohrer & M. Sherif (Eds.), *Social psychology at the crossroads*. New York: Harper, 1951.
13. Postman, L., Bruner, J., & McGinnies, E. Personal values as selective factors in perception. *J. abnorm. soc. Psychol.*, 1948, **43**, 142–154.
14. Postman, L., & Schneider, B. H. Personal values, visual recognition, and recall. *Psychol. Rev.*, 1951, **58**, 271–284.
15. Sanford, R. N. The effects of the abstinence from food upon the imaginal processes. A preliminary study. *J. Psychol.*, 1936, **2**, 129–136.
16. Sanford, R. N. The effects of the abstinence from food upon the imaginal processes. A further study. *J. soc. Psychol.*, 1937, **8**, 335–343.
17. Solomon, R. L., & Howes, D. H. Word frequency, personal values, and visual duration thresholds. *Psychol. Rev.*, 1951, **58**, 256-270.
18. Thorndike, E., & Lorge, I. *The teacher's word book*. New York: Teachers College, Columbia Univer., 1944.
19. Verplanck, W. S., & Hayes, J. R. Some relationships between "hunger" and "thirst" and "eating" and "drinking." *Amer. Psychologist*, 1951, **6**, 288. (Abstract)
20. Wispé, L. Physiological need, verbal frequency and word association. *J. abnorm. soc. Psychol.*, in press.

(Received Nov. 3, 1952)

Journal of Experimental Psychology
Vol 46, No. 1, 1953

AN ELECTROMYOGRAPHIC STUDY OF TENSION IN INTERRUPTED AND COMPLETED TASKS [1]

A. ARTHUR SMITH

Allan Memorial Institute of Psychiatry and McGill University

Lewinian theory (**3**) postulates the existence of "tension systems" associated with the carrying out of tasks. These normally disappear when task activity ends. If, however, the task is unfinished, the tensions will persist; and this will in general lead to differences when complete and incomplete activities are compared.

Most often, such comparisons have been made on a purely psychological level. The present investigation, using the technique of continuous electromyographic (EMG) recording, attempts: (*a*) to observe differences between completed and interrupted activities at a physiological level, that of muscle tension; and (*b*) to infer therefrom something about the nature of the tensions postulated by Lewin.

METHOD

Tasks.—Four tasks were administered, each of which called for the drawing of a simple geometrical figure. Figure 1 illustrates the patterns used. These were drawn separately with India ink on letter-size sheets of white paper, and presented to *Ss* as mirror-tracing tasks. The circumferences were 26 cm. for the smaller figures (always used as completed tasks) and 31 cm. for the larger (always interrupted). In the latter, *S* was stopped at a point 5 cm. from the goal, thus equalizing the work done on all tasks. The tasks were given in a different, random order to each *S*.

Apparatus.—Continuous tracings were taken of the bipolar EMG from five muscle groups. The recording apparatus was an Offner Type D electroencephalograph. The electrodes were attached to the forehead, the back of the neck (over trapezius), the chin (above and below the point, over muscles of the lower lip and tongue), and to the extensor surfaces of both forearms. Except for the neck, where chlorided silver electrodes of EEG type were used, the electrodes were small cubes of cellulose sponge, soaked in thin saline jelly.

The outputs of four of the main amplifiers were fed into direct-reading manually controlled integrators (**1**). The operating cycle of these instruments consisted of a 10-sec. integration period, followed by a 10-sec. interval for reading and discharge.

Subjects.—Fifteen *Ss* took part in the main experiment. Eight were college students (five female and three male); and seven (all male) were service personnel from an Air Force station. In addition, four Air Force men carried out a check experiment. The college group ranged in age from 18 to 21 years, the Air Force from 21 to 32.

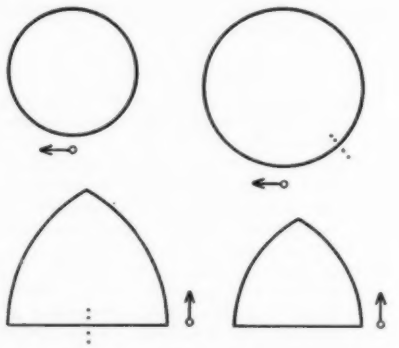

FIG. 1. Task materials. Drawing was carried out in the direction of the small arrows; interruption occurred in region of broken cross-lines (not present on material given to *Ss*).

[1] This research was performed under Contract No. DA-49-007-MD-70 between the U. S. Department of the Army, Office of the Surgeon General, and McGill University. The present report is based on a thesis submitted to the Faculty of Graduate Studies and Research, McGill University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy. An abridged version was read before the American Psychological Association, September 1952. Grateful acknowledgment is made of the interest and support of Dr. Robert B. Malmo in the execution of this research.

Questioning at the end of the session revealed that none of the *Ss* was familiar with previous work on interrupted tasks.

Procedure.—When the electrodes had been attached, each *S* was given a general outline of the way in which the experiment would be conducted. At this time, half of the *Ss* were told: "If for any reason you don't finish one of your tasks, don't worry about it." The other *Ss* were given no hint that they would be interrupted.

Each task was preceded by a 1-min. expectation period, during which *S* held his pencil on the start point of the drawing and waited without moving for *E*'s signal; and when activity had ceased—either because of interruption, or because the goal had been reached—*S* again remained motionless, holding his pencil on the point last reached for 1 min.

Interruption consisted in *E* saying quietly, "Stop. Keep your pencil on the paper." To preserve formal similarity, the same phrase was used when *S* had reached the goal.

Treatment of data.—The primary data consisted of: (*a*) the EMG record, in which the height of an individual spike corresponded to the instantaneous value of the potential difference between the relevant electrodes; and (*b*) the integrator readings, a set of numbers proportional to the average value of the potential difference during defined 10-sec. periods.

These data were reduced to microvolts of potential difference by reference to a standard or calibration signal, impressed on the recording and integrating circuits at the end of each session. The standard signal was a 45-cycle sinusoidal voltage, generated by a Hewlett-Packard audio-oscillator; its amplitude was controlled by a General Radio microvolter.

Two main classes of measurement proved feasible. To find the mean tension during one of the major time segments the integrator data were used, three converted integrator values being averaged for each minute of the test.

For short-term phenomena, a different procedure was necessary. This called for the measurement, on the EMG record, of the highest spike in each .2 sec. of certain designated 2-sec. periods. All measurements were made to the nearest half-millimeter. The ten measured spikes were averaged, and the average converted to a microvolt value.

Results

The integrators measured over-all tension during the 1-min. expectation- and postperiods, and during the task. Comparisons between interrupted and completed tasks on the basis of these data revealed no significant differences.

TABLE 1

Average Tension Change (in Microvolts) at End of Activity

Condition	Location of Muscles				
	Active Arm	Chin	Neck	Fore-head	Passive Arm
Completion	97.7	2.5	7.6	−0.6	10.0
Interruption	47.5	−3.6	1.3	−3.7	8.5
Difference	50.2	6.1	6.3	3.1	1.5
N*	15	12	15	15	14
t	†	3.65	2.34	1.53	0.58
p	.003	.01	.05	.10	.50

* Variation in N represents loss of data due to technical difficulties.

† The t test was not applied, due to a significant difference in variance estimates; the probability estimate was derived from the sign test.

The 2-sec. averages turned out to be more fruitful. They were designed to answer the question, "How much does tension decrease as soon as *S* stops drawing?" Two periods were chosen for comparison: the last 2 sec. of drawing, and the 2 sec. immediately following. The difference between these two periods gave the required measure. Table 1 presents the average change for each muscle group. Muscle tension is seen to fall more after completion than after interruption.

It is instructive to examine the actual tension levels from which these decreases were derived. Table 2 gives the values for the active arm; in addition, it indicates the tension level at various points during *S*'s progress toward the goal. To obtain these values, 2-sec. tension averages were obtained from measurements on the EMG trace. Because measurements along this record represent time rather than distance, the required points were located by assuming that *S* drew at a relatively constant rate throughout. Even if this assumption is not strictly accurate, the effect is still that of a smooth, progressive increase in ten-

TABLE 2

ACTIVE ARM TENSION AND DISTANCE FROM GOAL

Condition	Percentage of Task Completed					
	Start	25%	50%	84%	100%	Post*
Completion	89.9	151.5	161.0	172.4	196.6	98.9
Interruption	90.3	147.0	167.0	173.2	—	125.7

* The difference between the values in this column is significant at the 1% level of confidence.

sion, apparently as a function of distance from the goal. The shape of the tension curve appears most clearly in Fig. 2.

The data of Tables 1 and 2 are interconnected in a revealing fashion. When, for the active arm, the difference between completion and interruption is compared with the increase (averaged for all four tasks) from start to the 84% point, the rank-order correlation is .60. This indicates, in brief, that *S*s who showed the greatest rise in tension during drawing also showed the greatest tendency to maintain this tension when interrupted.

One point remains. Four *S*s were given the same test procedure, except that they were allowed to finish every task. For these *S*s, no difference was found between completed and "interrupted" tasks. In this context, "interrupted" refers to the larger figures, the ones which had been interrupted for the main group.

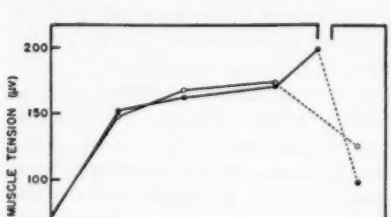

FIG. 2. Tension in active arm as a function of distance from goal

DISCUSSION

We have seen that muscle tension drops more after completion than after interruption. Is this due to interruption per se? Performance time is not, apparently, a factor; the mean times for completed and interrupted tasks are, respectively, 33½ and 34 sec.

More serious are the suggestions that the potentials do not decline after interruption either because that interruption was unexpected and hence "startling," or because *S* expected to go on with the task. It will be recalled that half of the *S*s were given no hint that they would not carry all tasks through to completion; the others were told not to worry if they did not finish. If, now, the observed differences were dependent on such factors, they should appear most clearly in the group which was not forewarned. Such is not the case. Where the difference between completion and interruption is (in active arm) about 50 microvolts for the group as a whole, it is only 25 microvolts for the group receiving no warning, and 75 microvolts for the rest of the *S*s.

It is now possible to consider the meaning of these findings. The hypothesis will be advanced that the observed increase in muscle tension

during drawing reflects the developing organization of a central neural process, which controls and "anticipates" the overt activity, and is yet relatively independent of immediate sensory stimulation.

Several facts point both to prior organization and to freedom from sensory control. In tracing the "triangular" figure, the first movement required is in exactly the opposite direction from that which seems correct to *S* looking in the mirror. And yet no *S* began in the wrong direction—surely a surprising result were there no previously organized process to anticipate the correct response.

Again, the reports made by various *S*s after the experiment can be cited. One *S* who had made a particularly accurate set of tracings remarked that he had "figured out ahead of time how to do it, and then sort of shut his eyes and did it." It will, of course, be remembered that before each task there was a 1-min. expectation period, which would allow ample time for such figuring out.

Another inference can be drawn. From the reports it would appear that carrying out the tasks was not just a matter of motor organization, but that an element of conceptual activity —of thinking—was involved. The suggestion is strengthened when the patterning of the muscle-tension changes is examined. We have seen (Table 1) that the effect under examination is confined either to muscles functionally involved in carrying out the movements, or to muscles which are associated most directly with speech, i.e., muscles of the tongue and lower lip. Now it is probably fair to assert that, at least in the human adult in our present culture, thinking involves a good measure of verbal activity. Consequently, specific activity in speech muscles may well imply the existence of something which could be labeled a "thought process."

In view of the picture that has been built up, what conditions are likely to obtain at the end of activity? When, in the completed tasks, the goal has been reached, the central process will by hypothesis already have run its course; cessation of activity will be accompanied by rapid relaxation of all associated muscle tensions. But, when activity is interrupted, the central process will tend to continue for some brief time, and thus maintain the task-associated tensions.

So far questions as to the nature of the central process and its mode of organization have been dodged. The problem is essentially the old one of "set," of "determining tendency," of directive and anticipatory processes in general. In a sense, organization is present from the start. But it is basically static: a pattern of lines in space; whereas the question here is one of the organization of processes in space and time.

Recently Hebb has developed ideas for handling such problems. In brief, the notions are those of "cell-assembly" and "phase sequence"; the cell-assembly being defined as "a diffuse structure . . . capable of acting briefly as a closed system, delivering facilitation to other such systems and usually having a specific motor facilitation. A series of such events constitutes a 'phase sequence' " (2, p. xix).

From this point of view, what can be said of the situation when an adult human *S* is confronted with the tasks here employed? Since these involved only simple movements of a pencil over a sheet of paper, it would seem that *S*'s problem was not how to execute any particular movement, but rather which of a number of familiar movements to carry out. In terms of the theory, *S* brought to the task an extensive set of previously developed cell-assemblies, which then became organized into a more comprehensive unit, the phase sequence.

Such organization could occur in part during the expectation period: the instructions, together with the pattern of sensory impulses resulting from *S*'s inspection of the task figure, would serve to facilitate selectively certain of the phase sequences, and to establish or improve connections between them. Such a process would be the neural correlate of "figuring out ahead of time what to do." The signal to start, then, triggers the phase sequence; and it in turn, by virtue of its specific motor facilitation, guides the overt activity.

This having begun, sensory feedback exercises a further selection among available neural structures. That is, at the beginning the phase sequence is thought of as a loosely knit aggregate, many of whose components could lead to alternate ways of doing the task; these variants compete for the final common path to the musculature. As the task proceeds, the error-correcting action of sensory feedback tends to the elimination of alternates, to leave (in the limit) a single closely knit and dominant sequence of assembly actions. It is thus that the central process becomes better organized; thus that the increase in muscular tension comes about.

Each cell-assembly is capable of self-maintained action, as is the phase sequence as a whole. Hence, if interruption does not too seriously disrupt the sequence, its activity will carry on for a short time. And this "carry over" will also be a function of the level of organization attained. That is, tensional differences between completion and interruption should be related to tensional increases during activity; this, we have seen, does occur.

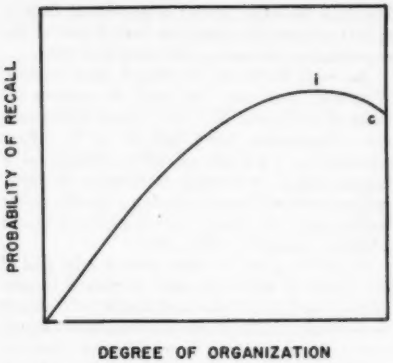

FIG. 3. Theoretical curve relating recall to organization of phase sequence. The points *i* and *c* represent the level of organization attained in interrupted and completed tasks, respectively.

Hebb (2, p. 143) has estimated the duration of such autonomous action at about half a second for a single cell-assembly, and a few seconds for a phase sequence. In the present instance, the fact that differences between completion and interruption are seen during the 2 sec. immediately following drawing, but are obscured when longer postperiods are compared, points to a similar order of magnitude.

This correspondence, while gratifying, is yet a source of theoretical difficulty. One of the implicit aims of the study was to throw light on the question of differential recall, as displayed in studies like that of Zeigarnik (4). A process which ceases well before the time when recall tests are administered can hardly be a sufficient answer.

One suggestion may be offered. The present hypothesis requires that phase sequences corresponding to finished tasks be better organized than those corresponding to unfinished ones. Now, Hebb has suggested that a well-organized phase sequence is not likely to be repeated in detail; it will tend to "short-circuit." Behavior will then tend to be "dominated by the thought process that is not fully organized" (2, pp. 227-230). The conception is, then, that when recall is asked for, phase sequences corresponding to completed activities have a greater probability of "short-circuiting," of running through too rapidly and in too abbreviated a form to control effectively the memory process.

There are obvious limits here. Too little organization can hardly be imagined to improve recall. A more satisfying hypothesis (illustrated in Fig. 3) is that subsequent recall of an activity is favored by increasing organization of the associated phase sequence, up to the point at which short-circuiting becomes likely. Beyond this point, the more highly organized process will become increasingly less likely to effect recall. This idea, it may be noted, is supported in some measure by Zeigarnik's finding that tasks interrupted when nearly complete were remembered better than when interrupted shortly after work had begun.

SUMMARY

The experiment sought to determine differences in muscle tension between complete and incomplete activities. Four mirror-tracing tasks were used, two of which were interrupted. A continuous EMG record was obtained from five muscle groups.

Activity produced a steady rise in tension; this was related to distance from the goal. Muscular tension in certain critical areas persisted for a time after interruption.

The findings were attributed to central neural processes capable of self-maintained activity. A relationship was postulated between the level of organization of such processes and the probability of subsequent recall.

REFERENCES

1. DAVIS, J. F. Integrator-counter for muscle potentials. Unpublished M. Eng. thesis, McGill Univer., 1949.
2. HEBB, D. O. *Organization of behavior.* New York: Wiley, 1949.
3. LEWIN, K. *A dynamic theory of personality.* (Trans. by D. K. Adams and K. E. Zener.) New York: McGraw-Hill, 1933.
4. ZEIGARNIK, B. Das Behalten erledigter und unerledigter Handlungen. *Psychol. Forsch.*, 1927, **9**, 1–85.

(Received October 27, 1952)

Journal of Experimental Psychology
Vol 46, No. 1, 1953

THE EFFECT OF DIFFICULTY OF TASK ON PROACTIVE FACILITATION AND INTERFERENCE [1]

ABRAM M. BARCH [2]

State University of Iowa

The purpose of this study was to investigate the kind and amount of transfer that would arise in successive practice on two motor tasks when the stimulus-response relations of the tasks were reversed and when the difficulty of the initial task was also varied.

The experimental design has relevance for two recent theoretical formulations of transfer in motor performance—one by Gagné (3) and the other by Lewis. The Gagné theory of motor learning, making use of the concepts of stimulus and response generalization, has had considerable success in predicting transfer effect. However, one finding which the theory has not handled is that transfer may be greater in going from a difficult to an easy task than from an easy to a difficult one (1, p. 729; 4, 10).

Lewis' theory.[3]—Lewis uses five constructs to explain not only the acquisition of proficiency on complex motor tasks but also the presence of simultaneous positive and negative transfer effects. These constructs are *dexterity*, *facilitation*, *skill*, *interference*, and *inhibition*. *Dexterity* consists of the inherent ability to perceive or comprehend task requirements and to execute the kinds of movements which are necessary for its accomplishment. *Facilitation* consists of acquired tendencies to make appropriate responses to the underlying, general, unchanging features of the task. *Skill* consists of acquired tendencies to make appropriate responses to the aspects of the task situation which change either randomly or systematically in clearly identifiable ways during the performance period. *Interference* consists of acquired or innate tendencies to make inappropriate responses in relation to *any* aspect of the task situation. *Inhibition* consists of acquired tendencies to inhibit or suppress in some way irrelevant or inappropriate responses.

Performance potential is regarded as being some function of all tendencies to make appropriate responses or inhibit inappropriate ones minus some function of all tendencies to make inappropriate responses. When *S*s practice on a task in the laboratory, a certain demand is made on their *dexterity*. Furthermore, past learning has resulted in response tendencies that may aid or hinder the acquisition of appropriate response tendencies toward the general features of the task. Only rarely will *S*s have prelaboratory training that will aid or hinder the acquisition of appropriate responses in relation to the specific aspects of the task.

When *S*s are required to practice on a second task that is a reversal of the initial task, certain response tendencies appropriate to the first task will become inappropriate. These tendencies will now increase the amount of *interference*. All of the *skill* becomes *interference*. If only a reversal is involved, little of the *facilitation* and *inhibition* is converted to *interference* (i.e., reversal of stimulus-response relations has little effect on the appropriate response tendencies to the general features of the task or on most of the irrelevant responses). If changes are introduced in the general features of the task at the same time that the stimulus-response relations are reversed, the resulting amount of *interference* may be greater than for a reversal alone because of a greater loss in *facilitation* and *inhibition*.

Thus, for Lewis, a complete reversal of stimulus-response relations alone may be less disruptive than a complete reversal plus a change in some of the underlying features of the task.

[1] This paper represents part of a dissertation submitted in 1952 to the Graduate School of the State University of Iowa in partial fulfillment of the requirements for the Ph.D. degree. The research was directed by Dr. Don Lewis and was supported in part by the United States Air Force under Contract No. AF 33(038)-13214, monitored by the Director of Research, Perceptual and Motor Skills Research Laboratory, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas. Permission is granted for reproduction, translation, publication, use, and disposal in whole and in part by or for the United States Government.

[2] Now with the Human Resources Research Office, George Washington University.

[3] This formulation is based on an unpublished paper by Lewis. However, earlier versions of Lewis' theory (2, 8), with supporting evidence, are available.

Similarly, a fractional reversal alone (reversal of a portion of the stimulus-response relations) may be less disruptive of performance than a fractional reversal plus a change in some of the underlying features of the task. For this to be so, the change in the underlying features of the task must be such as to result in a reduced level of appropriate response tendencies toward the general features of the task situation. The next section will show how changes in the stimuli and/or responses that increase the difficulty of the task are among the changes which can be regarded as reducing the level of appropriate response tendencies.

A special assumption.—The formulation by Lewis can incorporate the finding that transfer may be greater from a difficult task to an easy one than from an easy task to a difficult one by making the following assumption: The group that performs a more difficult task as its initial task acquires more *facilitation* and *inhibition* during this initial task than does a group performing with an easier task as its initial task. (The assumption is intended only for the case where the difficulty of the initial task is relatively independent of past training and largely dependent on the demands made on basic neuromuscular structure.)

This assumption is a reasonable one. The *S*s on an easy task can make maximum or near-maximum scores without learning, to any appreciable degree, to respond appropriately to such general features of the task as timing of work periods, general course of the target, etc., or without learning to inhibit such inappropriate tendencies as lags of attention, slowness of movement, bad posture, etc. On the other hand, most of the improvement displayed by *S*s practicing a difficult task may depend on learning just those aspects of the task.

Thus, a group that initially practiced on a difficult task and is tested on a reversed form of the difficult task would be expected to perform better on the reversed task than a group that first practiced on an easy form of the training task. The *skill* acquired prior to the reversal turns into *interference*. But the *facilitation* and *inhibition* acquired prior to the reversal remain relatively unaffected. Thus, if the group that practiced the difficult form of the task throughout had acquired a sufficiently greater amount of *facilitation* and *inhibition* than the other group prior to the reversal, it should perform better on the reversed form of the difficult task than the group with the initially easy task. This would hold for either a complete or a fractional reversal.

Problem

The purpose of this study, then, was to investigate: (*a*) the kind and amount of transfer in a two-hand coordination situation where the two tasks involved are complete reversals of each other (reversals of both controls) and where the change in task may also involve a change in the size of the target to alter the task difficulty; (*b*) the kind and amount of transfer in a two-hand coordination situation where the two tasks are fractional reversals of each other (reversal of one control) and where the change in task may also involve a change in the size of the target to alter the task difficulty.

Three tasks were employed. Task A is the standard coordinator task; Task B is the reversed form of Task A; Task C has only one of the two controls reversed.

The expectations were that: (*a*) The group trained with a small-sized target on Task A and tested with a small-sized target on Task B (complete reversal) would perform better on Task B than the groups trained with either a medium-sized or a large-sized target on Task A and tested with a small-sized target on Task B; (*b*) the group trained with a medium-sized target on Task A and tested with a small-sized target on Task B would perform better on Task B than the group trained with a large-sized target on Task A and tested with a small-sized target on Task B; (*c*) the group trained with a small-sized target on Task A and tested with a small-sized target on Task C (fractional reversal of A) would perform better on Task C than a group trained with a large-sized target on Task A and tested with a small-sized target on Task C.

Method

Subjects.—The *S*s were 79 female volunteers from the introductory course in psychology. Eight *S*s were eliminated because of experimental errors or mechanical malfunctions, and one *S* was eliminated because of failure to follow the instructions.

Apparatus.—The Modified Two-Hand Coordinator (9) was used. The *S* was required to manipulate two handles so as to keep a small

button, ½ in. in diameter, on top of a target that moved in an irregular circular pattern in the horizontal plane at the rate of 1 rpm. Only the 30-sec. period during which the target was nearest to *S* was used. The same pattern of movement was presented on every trial.

The right-hand handle turned in the vertical plane which is parallel to the frontal body plane and moved the button toward the right or left. The left-hand handle turned in the vertical plane that is perpendicular to the frontal body plane and moved the button toward or away from *S*.

One complete revolution of either handle moved the button ½ in. The targets were ¼ in., ½ in., and ¾ in. in diameter.

Measures of both correct and incorrect responses were available. Time on target was recorded to the nearest .01 sec. The amount of time that each control was moved so as to increase the distance between the button and the target (persistence of moving error) and the amount of time that each control was stationary even though the button was not on target (stationary error) were obtained.

Tasks.—In Task A, the Standard Task, clockwise turning of the right handle moved the button to the right and counterclockwise turning moved it to the left. Clockwise turning of the left handle moved the button toward *S*, and counterclockwise turning moved the button away from *S*. In Task B, the Reversed Task, the handles had to be moved in exactly the opposite direction from Task A. In Task C movements of the button produced by the right handle were the same as in Task A while movements of the button produced by the left handle were the same as in Task B.

Experimental design.—The *S*s were divided into five groups of 14 each. On Day 1 all groups practiced 25 trials of Task A. Groups S–S and PS–S used the ¼-in. or small-sized target; Group M–S used the ½-in. or medium-sized target; Groups L–S and PL–S used the ¾-in. or large-sized target.

On Day 2 all groups first had five trials on Task A using the same target as on Day 1. Groups L–S, M–S, and S–S then had 20 trials on Task B with the small-sized target. Groups PL–S and PS–S practiced 20 trials of Task C with the small-sized target.

Procedure.—All *S*s were given complete instructions as to the nature of Task A at the beginning of the experiment. A warning light came on 5 sec. before each trial began. Trials were 30 sec. in length; intertrial rests were also 30-sec. A 3-min. rest was given after every block of five trials. The *S*s were required to return the button to a fixed starting point at the end of every trial and were on target at the beginning of every trial. No information as to recording of errors was given.

After the first five trials of Day 2, instructions were given as to the nature of the alteration in task and attention was called to the change in target in the groups with a change.

Results

No general test for initial comparability was made since the various groups during Task A practiced on different sizes of target. It was assumed that a procedure of random assignment of *S*s to the various groups gave comparability. Support is provided for this assumption by the near identity of the mean scores for the first ten trials of Task A for the two small-sized target groups (7.87 and 7.78 sec.) and for the two large-sized target groups (22.38 and 22.40 sec.).

Difficulty of task.—The groups were compared to see if difference in size of target yielded tasks of differing difficulty. Groups L–S and PL–S were combined, as were Groups S–S and PS–S. Because a number of *S*s in the large-target group achieved the maximum possible score quite early in practice, thus restricting the range and introducing skewness, the later trials of Task A were not suitable for a comparison of difficulty. For this reason, the measure used was the mean score made by each *S* on the first ten trials of Task A.

The analysis of variance for the difference in mean time on target during the first ten trials of Task A for the various target sizes was highly significant. The *F* obtained was 136.63 for 2 and 67 *df*. The means of time on target for the three groups were 22.39, 14.95, and 7.83 sec. The large target made the task significantly easier than the medium- or the small-sized target ($p = .001$); the medium-sized target made it significantly easier than the small-sized one ($p = .001$). Thus, use of different sizes of target did lead to tasks of differing difficulty.

Reversal of both controls.—Figure 1 presents performance curves for Groups L–S, M–S, and S–S, based on

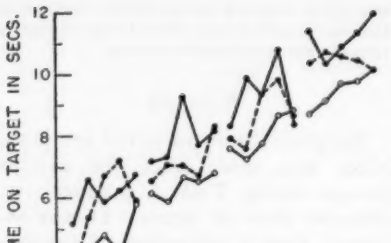

FIG. 1. Mean time on target, on the trials of Task B for Groups L–S, M–S, and S–S

the means of time on target for the trials of Task B. The general tendency is for the performance of Group S–S to be superior to the performance of Group M–S, and for the performance of Group M–S to be superior to the performance of Group L–S. There is some overlapping of the curves and some irregularities in the trend.

The trend analysis for mean time on target for all trials of Task B did not indicate a significant difference among the groups. (See Table 1 for form of analysis.) The F obtained for the effect of target size was 1.23 for 2 and 36 df. The interaction between size of target and stage of practice on the test task was also not significant (F = .24 for 6 and 117 df). The F for stage of practice was highly significant (p = .001) for this comparison and for all similar ones, merely indicating that performance on the test task improves with practice. The measures for each S were mean scores for blocks of Trials 1–5, 6–10, 11–15, 16–20 of Task B.

TABLE 1

A TREND ANALYSIS OF TIME ON TARGET DURING TASK C

Source	df	Mean Square	F
Target size	1	311.49	6.30*
Error (between)	26	49.48	
Stage	3	195.16	66.84**
Stage × Size	3	3.98	1.36
Error (within)	78	2.92	
Total	111		

* Significant at 5% level.
** Significant at 1% level.

On the other hand, the F for mean time on target for the first trial of Task B for Groups L–S, M–S, and S–S was significant at the 1% level (F = 6.87 for 2 and 39 df). The means of time on target were 2.26, 3.50, and 5.20 sec., respectively.

The performance of Group S–S was significantly better than that of Group M–S (p = .05) and of Group L–S (p = .01). Group M–S was not significantly different from Group L–S.

Total error for a given S for a given trial was obtained by summing right and left stationary errors and right and left movement errors. A trend analysis for total error for all trials of Task B for these three groups was not significant. The F obtained for the effect of target size was 2.77 for 2 and 39 df. The interaction between size of target and stage was also not significant (F = .48 for 6 and 117 df). The

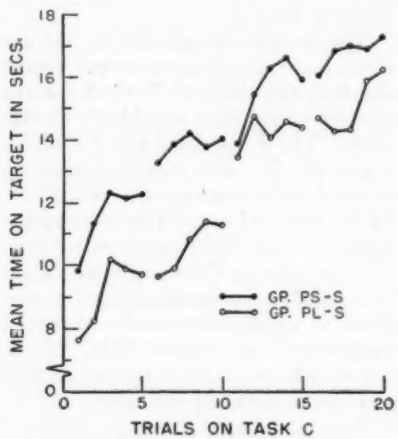

FIG. 2. Mean time on target, in seconds, on the trials of Task C for Groups PL–S and PS–S

measures used for each *S* were her mean total errors for blocks of Trials 1–5, 6–10, 11–15, 16–20 of Task B.

An analysis of variance for total error on the first trial of Task B for these three groups was also not significant. The *F* was 2.77 for 2 and 39 *df*.

Reversal of both controls plus a change in target size yielded significantly worse performance than reversal of both controls without a change in target size, but only with respect to time on target during the first trial of reversed practice.

Fractional reversal.—Figure 2 presents performance curves based on the means of time on target for Groups PL–S and PS–S during Task C. As can be seen in Table 1, a trend analysis for mean time on target for all trials of Task C was significant at the 5% level. The *F* for the effect of target size was 6.30 for 1 and 26 *df*. The interaction between size and stage was not significant ($F = 1.36$ for 3 and 78 *df*). As in the analysis of the groups with both controls reversed, the means of blocks of five trials were used as the individual measures.

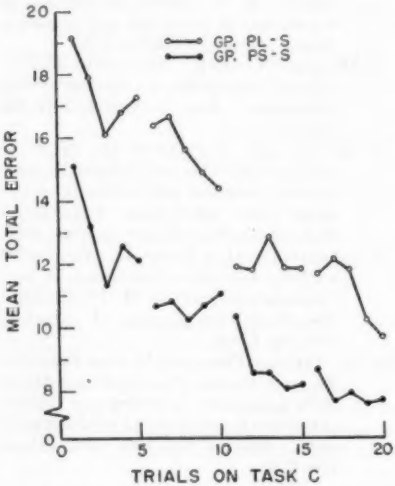

FIG. 3. Mean total error on the trials of Task C for Groups PL–S and PS–S

Figure 3 depicts the means of total error during Task C for Groups PL–S and PS–S. Again a clear difference between the two groups is displayed. The *F* for the effect of target size was 4.74 for 1 and 26 *df* and was significant at the 5% level. The interaction between size and stage was again not significant ($F = 1.45$ for 3 and 78 *df*). These measures were also means of blocks of five trials.

Thus, when only one control was reversed, the reversal plus a change in target size led to significantly poorer performance than did reversal alone. The difference in performance level was apparent in both time on target and the total error scores for the entire 20 trials of the reversed task.

DISCUSSION

In general the results supported the predictions that a group undergoing a reversal of stimulus-response relations plus an increase in the difficulty of the task would display more negative transfer or less positive transfer than a group undergoing only a reversal of stimulus-response relations. The relationship, however, was much more pronounced for a fractional reversal than for a complete reversal. Although the performance of Group S–S was significantly better than the performance of Groups M–S and L–S only with respect to mean time-on-target scores for the first trial of Task B, the performance of Group PS–S was significantly better than the performance of Group PL–S throughout Task C both in mean time on target and mean total error.

This suggests that the amount of *facilitation* and *inhibition* indicated in the overt responses may not be independent of the amount of *interference* present, i.e., the greater the amount of *interference*, the less the external indication of the amount of *facilitation* present. No provision for such an interaction exists in the present formulation. It has been assumed that *facilitation* and *interference* are independently measurable.

Although the work and rest arrangements

used throughout the experiment were considerably more distributed than has been the arrangement in past studies with the Two-Hand Coordinator (5, 7, 9), there seems, nevertheless, to be some indication in the performance curves of Fig. 1 of the accumulation of work decrement. This is shown by the fact that groups performing at a relatively high level for a given stage of practice undergo quite abrupt and sizable losses in performance level, especially toward the end of a block of trials. The possible presence of work inhibition in this study, however, does not bias the result in favor of the stated expectations since the effect of work inhibition would be to make Group S–S (and Group PS–S) less superior to the other groups than would be the case if no work decrement were present.

These results do not necessarily invalidate Gagné's general approach to the problem of transfer in motor learning. They do point to the need for the inclusion of a concept of task difficulty in his present theoretical structure for situations involving reversal of stimulus-response relations as well as for situations that do not involve such reversals.

Moreover, the necessity still remains for a fuller recognition of the experimental fact that is the cornerstone of Lewis' formulation; namely, that both negative and positive transfer effects can and often do appear simultaneously. For example, although mean time on target usually increases from the first trial of the original task to the first trial of a transfer task, the mean errors likewise increase.

Summary

The purpose of the study was to compare the amount of transfer between motor tasks that were reversals of each other when the relative difficulty of the initial task was manipulated by changing the size of the target.

Five groups of 14 female college students served as *Ss*. During the initial task on the Two-Hand Coordinator, two of the groups practiced with a large-sized (⅝ in.) target, two groups practiced with a small-sized (⅛ in.) target, and one group practiced with a medium-sized (¼ in.) target. After 30 trials of initial practice both controls were reversed (Task B) for a large-, medium-, and small-sized target group, and one control was reversed (Task C) for a large- and small-sized target group. All groups practiced with the small-sized (⅛ in.) target on the transfer tasks.

Use of different target sizes during initial practice led to tasks of differing difficulty. On the first trial of Task B, the group that had the small target on Task A was significantly superior to the group that had larger targets on Task A. Similarly, for the entire 20 trials of Task C, the group that had the small target on Task A was significantly superior to the group that had the large target.

The results warrant the following conclusion: As much or greater transfer is obtained between motor tasks that are stimulus-response reversals of each other than between motor tasks that are not only reversals of each other but also differ in difficulty, with the initial task being easier than the final task. This was true for a complete reversal as well as for a fractional reversal of stimulus-response relations.

References

1. Baker, K. E., Wylie, R. C., & Gagné, R. M. Transfer of training to a motor skill as a function of variation in rate of response. *J. exp. Psychol.*, 1950, **40**, 721–732.
2. Dusek, R. Facilitation and interference in the performance of simplified tasks on the modified Mashburn apparatus. Unpublished Ph.D. thesis, State Univer. of Iowa, 1951.
3. Gagné, R. M., Baker, K. E., & Foster, H. On the relation between similarity and transfer of training in the learning of discriminative motor tasks. *Psychol. Rev.*, 1950, **57**, 67–79.
4. Gibb, C. B. Transfer of training and skill assumptions in tracking tasks. *Quart. J. exp. Psychol.*, 1951, **3**, 99–110.
5. Lewis, D., Smith, P. N., & McAllister, D. E. Retroactive facilitation and interference in performance on the Modified Two-hand Coordinator. *J. exp. Psychol.*, 1952, **44**, 44–50.
6. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1953.
7. McAllister, D. E. Retroactive facilitation and interference as a function of level of learning. *Amer. J. Psychol.*, 1952, **65**, 218–232.
8. McFann, H. A study of the effects of response alteration and different instructions on proactive and retroactive facilitation and interference. Unpublished Ph.D. thesis, State Univer. of Iowa, 1952.
9. Shephard, A. H., & Lewis, D. Devices for studying associative interference in psychomotor performance: II. The Modified Two-Hand Coordinator. *J. Psychol.*, 1950, **29**, 53–66.
10. Air Training Command Human Resources Research Center, Perceptual and Motor Skills Laboratory. Learning and transfer of training in two forms of rudder control tasks. *Research Note* Lab. 50–1, February, 1950.

(Received November 8, 1952)

Journal of Experimental Psychology
Vol 46, No. 1, 1953

ACQUISITION OF TWO LEVER-POSITIONING RESPONSES PRACTICED OVER SEVERAL PERIODS OF ALTERNATION[1]

EDWARD A. BILODEAU

Perceptual and Motor Skills Research Laboratory, Human Resources Research Center

In psychomotor tasks *S* must receive some kind of knowledge of results in order to reduce his response error. The knowledge of results provided is often some measure of the difference between the response required and the response performed. In a simple situation requiring a single response of given amplitude, on the basis of the magnitude and direction of error signaled for the preceding response, *S* is able to adjust his next response to reduce the error. Over successive trials, response strength is built up, *S* responds with considerable accuracy, and receives signals of slight error.

Relatively uninvestigated, however, is the task situation where over successive blocks of trials two responses are required alternately, that is R_1, R_2, R_1, R_2, etc., where the particular response required at any moment is cued (signaled to *S*) by means of the response error associated with the immediately preceding trial. In such a case the error magnitude is determined by whether R_1 or R_2 was appropriate. If *S* assumes that the error is appropriate to R_1, but response R_2 is actually the reference, then *S* has confused the meaning of the error signal. What *S* attributes as an error deviation from R_1 is an error deviation from R_2. If *S*, nevertheless, next corrects the response in accord with the signal of previous error, his responses will show a pattern of alternation. Were *S*, moreover, to respond imperfectly (a motor error), the confusion between the selection of the appropriate response (R_1 or R_2) and inaccuracy of responding would become compounded. In other words, it may be difficult for *S* to show proper response alternation when the error cue represents a combination of alternate reference and previous response error. This situation is no doubt made more difficult by decreasing the number of trials constituting a block during which either response R_1 or R_2 is correct. When one response is correct for many trials, *S* can assess directly the error attributable to the last response and can be more certain which of the two responses is appropriate. Were *S* required to respond with a single-trial alternation pattern, the task would become very difficult.

The alternation of required responses as indicated above is related to Lewis' studies of retroactive inhibition where a second task is interpolated between single test and retest periods (3). In that repeated alternations are required, it is similar to that of Harlow's learning-set paradigm where *S*s have been found learning to learn (2). Typically, however, Harlow's procedure, as opposed to the present technique, never repeats the same problem and always equates successive problems for difficulty. The learning-set results have shown that later problems are mastered more quickly than the

[1] The experimental work for this study was performed as part of the United States Air Force Human Resources Research and Development Program. The opinions or conclusions contained in this report are those of the author. They are not to be construed as necessarily reflecting the views or indorsement of the Department of the Air Force.

earlier. For the problems used by Harlow, however, no response error complicates the cue received by *S*. There is no degree of incorrectness; for example, either alternative one or two is correct. A failure to receive the goal means the remaining alternative is (or was) proper.

The research questions (transfer, positive or negative) associated with what happens to responses as a function of the number of previous shifts in required responses can be readily extended to human motor skills by regulating *S*s' responses through direct experimental control (variation) of informative feedback. Previous motor-skills research with successive alternation schedules has suggested no improvement in the skill level of either of two responses as a function of number of previous shifts (1). Actually, the present experiment is the second attempting to relate the acquisition of two or more habits learned in alternation to controlled variations of informative feedback. The purposes of the present study are to check the reliability of the previous results and to investigate alternation behavior where the number of trials per alternation block is the variable.

Method

Subjects.—A total of 246 basic trainee airmen at Lackland Air Force Base without previous experience with the apparatus were used as *S*s.

Apparatus.—A schematic side view of the apparatus is given in Fig. 1. The stimulus panel is shown at the left—a vertical bank of 44 ½-in. red jeweled lamps. These lamps can be arbitrarily designated as 1 through 44 with No. 1 at the base of the column. A second column of 44 green lamps parallel to the first served as the response bank.

The response lever is also depicted in Fig. 1. The lever was so arranged that when grasped by means of the rubber handle at the top, it could be rotated through a total arc of 42°. The *S* sat in a chair before the apparatus in such a position that his hand could be placed upon the lever handle when his arm was nearly fully extended.

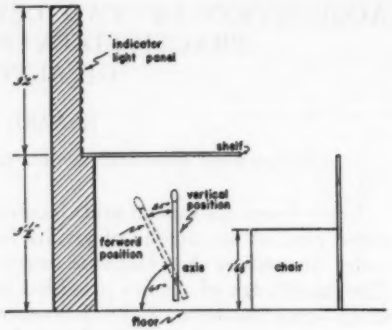

Fig. 1. Schematic drawing of side view of the apparatus. The stimulus panel is shown on the left side and the response lever is located beneath the shelf. *S* is seated such that the shelf obscures the lever. The farther the lever is pulled away from the forward position, the higher a light will appear in a vertical bank of green jewel lamps.

A wooden shelf at high chest level prevented *S* from viewing any position of the lever.

Procedure.—The *S* was required to position the lever in order to obtain a match of two lights upon the panel. He was told that pulling the lever back towards the body would "move" the response light up the bank of green lamps. (A lamp which is lighted will be referred to as a light.) A fixed red light (No. 26 of the stimulus bank) was pointed out, and *S* was instructed to pull the lever back to the point where he *thought* this light would be matched. The response light, however, was to appear 3 sec. after the completion of the response. Any discrepancy between response light and stimulus light (the error) could be used to reduce the magnitude of error on the subsequent trial. The *S*s were never told, however, that at times one lever setting would achieve the match and that at other times a different placement was required to achieve the same match. Throughout the experiment the same stimulus lamp remained lighted, and the response lever was automatically repositioned in the forward position at the beginning of each trial. Between responses *S*'s arm remained at his side.

Experimental design.—The *S*s were assigned unsystematically into five experimental groups and one control group of 41 *S*s each. They were run one at a time and the groups were run in a simple counterbalanced order. A total of 48 trials (one response per trial) was administered to all *S*s. The trial cycle was 20 sec., 5 sec. of which was used for a ready interval, 4 or 5 sec.

for the response, 3 sec. for the delay in information, and the remaining seconds for *S*'s evaluation of the response information provided.

For the experimental groups the constant +4 was added to the true setting and this total was reported on the trials of alternate blocks. The experimental groups differed solely in the number of trials constituting these blocks, i.e., either 1, 2, 4, 6, or 8 trials for the block definition. For example, Group 8 was given its first eight trials with true information, then a second block of eight trials during which the true settings were modified by adding the constant +4.[2] The blocks of trials during which the information was unmodified are called "U" periods; blocks of trials during which information was modified are called "M" periods. The control group (0) was given U periods throughout; informative feedback was never modified in any way during the 48 trials.

Scoring.—The response lamp corresponding to the true lever position was the score recorded. Thus, Response 26 means that the lever was pulled to the position corresponding to Lamp 26. During U periods the response measure is identical with the reported score, but during M periods the response measure differs from the reported score. The significance of the conversion of true setting to reported setting can be illustrated by means of a hypothetical *S*. If *S* has responded with 26 during a U period, he has achieved a match, Lamp 26 is lighted, and 26 is recorded. In other words, he has achieved the goal. But if *S* next responds with 26 during an M period, Response Lamp 26 + 4 or 30 is lighted while 26 is recorded. Assuming that *S* then realizes that he is four units too high and that this discrepancy can be corrected, the next response would be to pull the lever less far and to the lever position corresponding to Lamp 22. Provided the trial is an M, a match is reported to *S* by lighting Lamp 26, and Response 22 is recorded by *E*. Therefore, the requirements for all experimental groups were that they respond with 26 during U periods but with 22 during M periods.

[2] The design called for an equal number of responses within each period. Thus, for example, Group 8 should have eight responses within each of six periods ($8 \times 6 = 48$). Unfortunately, because of a misunderstanding, the experiment was completed so that the first period of each group contained one response more than called for and each final period, one less. Thus, the responses of Group 8 are grouped 9-8-8-8-8-7, and the responses of the remaining groups are similarly arranged. This state of affairs is reflected in the figures which follow.

Results and Discussion

The treatment of data consisted in calculating the average response and *SD* on each of the 48 trials for each of the six groups.[3] The trial-by-trial means are presented in Fig. 2. The figure has been constructed by plotting the groups over the same abscissa, but each group has its own ordinate, and a dashed horizontal line has been drawn opposite Response Lamp 26, the optimal response during U or odd-numbered periods.

The means of Group 0 show quite clearly that learning to position the lever to 26 is an easy task. Though the variability from trial to trial should be noted, the means of Group 0 do not differ materially from 26 and various consecutive points show consistency of response.

The interpretation of the results can be facilitated by using the means of Group 8 as an example. It will be recalled that the first eight responses for Group 8 were followed by information appropriate to strengthening Response 26. This meant, however, that nine U responses were appropriate. Thus, the first nine responses are shown connected, and subsequent responses are grouped according to U or M. Group 8 shows quite graphically that when 26 is appropriate, the responses tend to be clustered about or tend towards 26, but that when 22 is required, there is a trend towards 22. Thus, the trend of the means is entirely consistent with the changes in information. There is no tendency for *S*s to be progressively more successful in reaching the appropriate response (26 or 22) as a function of the number of previous shifts in required response.

[3] An inspection of the *SD*'s showed group variability to decrease with successive trials. The measures gave no reason to suspect real between-group differences.

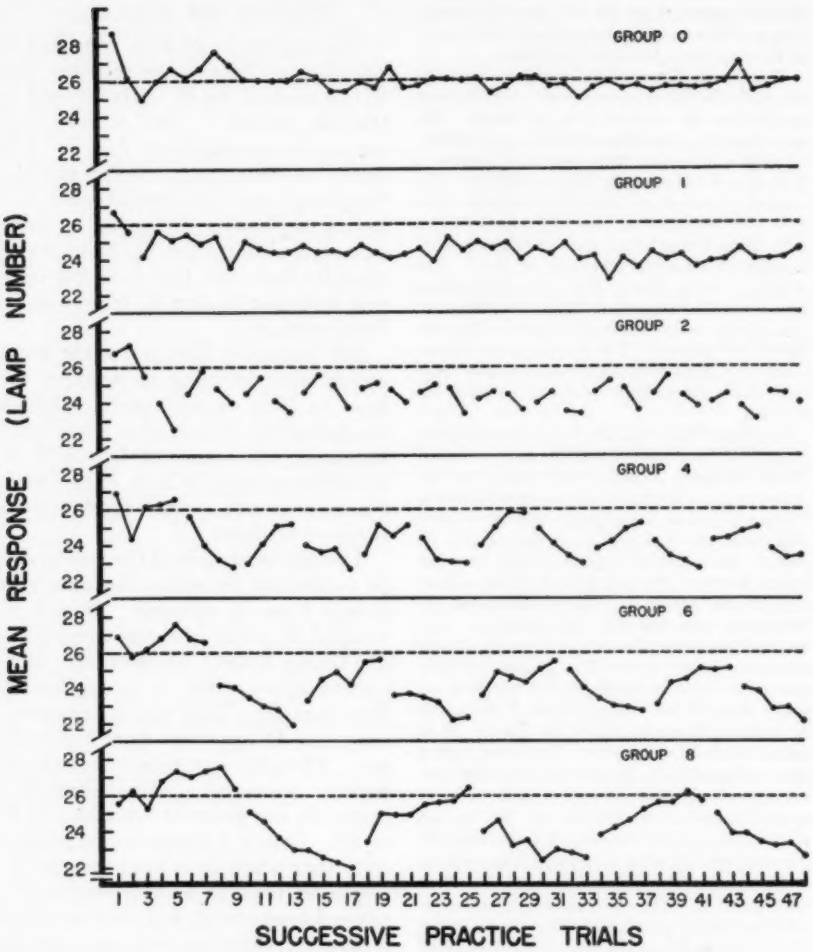

FIG. 2. Mean response is plotted as a function of successive practice trials for each group. Each group is plotted over the same abscissa, but with different ordinate. The dashed horizontal line represents optimal responding during odd-numbered periods (U), except for Group 0, where it represents optimal responding for all periods.

The same basic interpretation is applicable to Groups 6, 4, and 2. That is, whenever Response 26 is appropriate, the means in Fig. 2 trend in that direction, but whenever 22 is the correct response, the means trend toward 22. Also, for all these groups there is no greater tendency to be more correct in mean response during terminal blocks than during initial blocks. As a matter of fact, if the terminal trials of successive blocks are considered, Groups 6, 4, and 2 appear to be getting progressively worse and a compromise

solution is suggested. That is, relatively late in practice Response 24, midway between 26 and 22, appears to be the approximate mean value for either U or M blocks. This indicates that response resolution consists, progressively, in responding on each practice trial midway between the two required alternatives regardless of the response which the information of the previous trial implies. Examination of the response distributions shows them to be unimodal, thus precluding the possibility that the compromise phenomenon is an artifact of averaging.

If the analysis of alternating and compromising above is valid, Group 1, which is given single alternation of information, might be expected (*a*) to alternate responses such that alternate responses approach either 26 or 22 and (*b*) to trend toward 24. Indeed, most of the means of Group 1 show the single alternation pattern although there are four or five segments where the alternation pattern is broken. It is quite clear, however, that there is a drift away from Response 26 and, progressively and consistently, toward Response 24; evidently this group, too, had effected a unique solution—one which is midway between the two alternatives. Associated with the compromise response should be a tendency for a decrease in magnitude of response alternation. This could be called a damping phenomenon. Damping is not inappropriate for to the extent that Group 1 alternates its responses, as in Fig. 2, any given response is in opposition to the information associated with that trial.

Accuracy at lever settings during

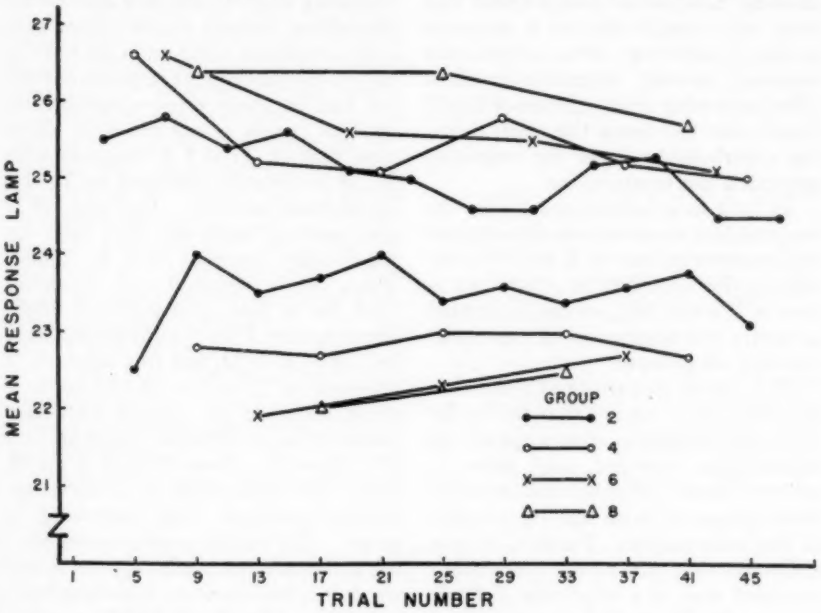

FIG. 3. Mean response for the terminal trial of each U or M period is plotted against trial number. The upper set of four curves represents performance when Response 26 is appropriate (U); the bottom four curves represent performance when Response 22 is appropriate (M).

the terminal trial of successive U or M periods is represented by the means of Fig. 3. The upper set of curves represents four of the experimental groups in the last trial of successive U periods, whereas the bottom set of curves represents mean terminal response over successive M periods for the same four groups. Because the former lines tend to have a negative slope, the groups, over successive U periods, were progressively less proficient at approximating Response 26. The magnitude of response over successive M periods, however, tends to increase or to show the opposite slope. What is indicated is that these groups are becoming less proficient over successive M periods at responding with 22. The interpretation of both groupings of data appears to be consistent with the hypothesis that all groups are drifting towards a compromise and that this compromise is a response midway between those which are required during alternate periods. The patterning of the groups of Fig. 3 shows that the fewer the trials defining a period, the faster the responses approach the compromise.

To lead to a definite conclusion the analysis just presented should be based on greater numbers of U and M periods; further study where greater numbers of U's and M's are used is needed to verify the suggestion of a compromise for all groups.

As in the earlier study (1), the data as a whole have shown that the *S*s did not take maximum advantage of the information provided, and only, in general, showed response trends which were consonant with the significance of the information. Further, it was indicated that when the information provided was of a relatively difficult nature, the *S*s tended, progressively, to approach faster a compromise between the two required responses, rather than to learn progressively to make better use of the information.

The trend of the means toward Response 24, midway between the two responses required during alternate periods, regardless of the specific period obtaining, has been called a compromise solution. It is a compromise in the sense that such a mode of attack results in (*a*) a manifest deviation from a perfect response *on each trial* and (*b*) a deviation which is smaller than the deviation resulting from the selection of the *wrong* alternative. Compromising in this fashion is most reasonable for groups where the alternation period is short, for it is here that the cue-response system is most complex and the probability of error greatest. The greater task complexity for the single alternation group, for example, can be illustrated in the following way: To simplify matters for exposition, assume at the outset that any *S* responds to his cue such that he can on the subsequent response correct for the previous error—*provided the external system is not changed.* Suppose that on Trial 1 *S* responds with 28, 28 is reported, and that on Trial 2 he responds with 26. However, 26 is not reported since this is a trial for which the constant +4 is added. Thus, 30 is reported. *S*, assuming that he is four positions too high, decreases the Trial 3 response by four, i.e., $26 - 4 = 22$, but this response is reported as $22 + 0 = 22$ and he, in a sense, expected 26. *S* next adds four positions to his (Trial 4) response, i.e., $22 + 4 = 26$. Since Trial 4 is an M trial, once again a signal of deviation of four positions from matching is given. If *S* were to continue responding in this manner, each response would result in a four-position deviation from matching. To obtain 100% matches requires 26's on odd-numbered trials and 22's on even-numbered trials.

But, if *S* responds on the basis of an evaluation of the signals of the *previous* trial, no matches can be obtained. Rather, the important cue is within Trial N-1, where N represents the previous trial. The response alternation of Group 1 showed rather clearly that responding was not cued by Trial N-1, but by Trial N. Such response alternation is always followed by reports of relatively large errors of responding. Thus, this may act as an extinction series, inasmuch as all response 22's and 26's cued by Trial N are in gross error. If, as a consequence, these two responses are avoided (damped), then the response midway between them comes to be more dominant.

Summary and Conclusions

A number of *Ss* practiced positioning a lever in order to obtain a match on a visual display panel. During odd-numbered blocks of trials one particular setting of the lever achieved the match, whereas during even-numbered blocks another setting of the lever achieved the same match. The *Ss* were never instructed about this state of affairs, but had to determine their course of action from information (discrepancy from match) supplied from the display panel. The groups differed solely in the number of trials constituting a block.

The data confirmed the conclusions of a previous investigation with respect to trends in response accuracy within and between successive information periods.

From the present experiment it is possible to conclude that regardless of the number of trials constituting a block or the number of previous shifts in required settings, accuracy of responding over successive periods failed to improve. Actually, some of the means for the first trial on which a shift in response setting was appropriate showed increasing departure from the proper setting with successive shifts, a phenomenon similar to interference.

More important was the finding of a progressive damping of response alternation such that *S* shifted toward a compromise response midway between the two required. Such a response assured *S* of some amount of error on each trial, but an error not as large as if the wrong response were selected. The compromise solution was more strongly and more quickly shown by the groups with the fewest trials per alternation period. The problem confronting the *Ss* of these groups is undoubtedly more difficult than were the periods of longer duration.

References

1. Bilodeau, E. A. Some effects of modification of information about a previous response upon the acquisition of two lever positioning habits. San Antonio, Tex.: Human Resources Research Center, Lackland Air Force Base, January 1952. (*Research Bulletin* 52-1.)
2. Harlow, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51–65.
3. Lewis, D. Devices for studying interference in psychomotor performance: III. The double-disk pursuit apparatus. *J. gen. Psychol.*, 1951, **45**, 123–132.

(Received November 14, 1952)

Journal of Experimental Psychology
Vol 46, No. 1, 1953

SOME EFFECTS OF MANIFEST ANXIETY ON MENTAL SET [1]

IRVING MALTZMAN, JACK FOX, AND LLOYD MORRISETT, JR.

University of California at Los Angeles

In a previous paper Maltzman and Morrisett (**4**) outlined a behavioristic account of directed thinking. It was assumed that the principles operating in discrimination and selective learning are largely responsible for the direction of thinking. It was further assumed that the aspect of directed thinking commonly called "mental set" is descriptive of the effects of the acquisition of a differential amount of habit strength by a particular class of stimuli for the elicitation of a particular class of responses. It involves a change in both the initial response and stimulus hierarchies. Another study by Maltzman and Morrisett (**5**) has shown that the direction of thinking can be influenced by factors other than habit strength. The experiments reported here are a further study of the effects of nonassociative factors on the direction of thinking. In the present case the effects are those of manifest anxiety.

Effective reaction potential, hence the direction of thinking, is a multiplicative function of habit strength, effective drive state, and other variables. The effective drive state is a function of the needs present in the organism at a given moment. Relevant needs that are reduced by the acts under investigation and irrelevant needs that are not so reduced may contribute to the effective drive state. An increase in reaction potential may therefore occur following an increase in habit strength with drive and other known variables constant, or following an increase in drive with habit strength constant.

The results of a recent experiment by Cowen (**1**) in which increased mental set was obtained with the introduction of stress may be accounted for in terms of the above theory. This investigator found that the introduction of various kinds of stress increased the number of set solutions in Luchins' water-jar problems. Training was with three jars in which the quantity to be obtained could be discovered by application of the principle B-A-2C. Crucial problems then followed that were solvable by the above indirect method or a more direct method such as A-C. An extinction problem presented could be solved only by the direct method. It was found that the introduction of stress increased the number of indirect solutions to the crucial problems and tended to make the extinction problem unsolvable.

According to the present account the experimental stress induced an irrelevant need which produced increments in the effective drive state multiplying all habits present at the moment. A greater absolute increase in the effective reaction potential of the dominant class of responses would therefore result. In the above experiment the dominant class of responses produced the indirect method of solution. An increased effective drive state therefore produced a further increase in the tendency for this method of solution to occur. However, if the dominant class of responses in a problem-solving situation produced a correct or a direct solution, an increase in irrelevant need would produce an increase in these solutions, within limits.

Employing the water-jar problem and anagrams as the problem materials, and manifest anxiety as the irrelevant need, the two experiments reported here test the following implications of behavior theory for the phenomena described as directed thinking. (*a*) If the dominant "set" is congruent with an indirect and incorrect solution, there will be an inverse relationship between strength of anxiety and number of direct and correct solutions

[1] This study was aided by a grant from the Committee on Research, Graduate Council, University of California.

in the water-jar problem situation. (*b*) If the dominant "set" is congruent with the correct solutions to a class of anagrams, fewer errors will occur under high than under low anxiety.

Method

Experiment I

Subjects.—Sixty student volunteers from undergraduate psychology classes were tested individually with the water-jar problem. An anxiety scale was administered to 48 of these *Ss* as a group during their regular class period. The data to be presented were obtained from the 48 *Ss* that received the water-jar problem and the anxiety scale. The restricted *N* was due to absences from the class.

Procedure.—Ten water-jar problems were presented to each *S*. The first five training problems were all solvable by the indirect method. The next two problems could be solved either by the indirect method or a direct method. The eighth or extinction problem could be solved only by the direct method. The last two problems were similar to problems six and seven in that they could be solved either by the indirect or the direct method. An interval of 2.5 min. was permitted for each problem. Incorrect solutions were not corrected and the correct solution was not given if *S* failed to solve a problem within the allotted time.

Each *S* received two practice problems before the start of the experiment. The first involved only two jars, while the second problem involved three jars and was solvable by the indirect method. The problems and general procedure were obtained from Luchins' original study (**3**).

Approximately one month after completion of the mental-set experiment the anxiety scale developed in the psychological laboratories of the State University of Iowa (**6, 8, 9**) was administered to the psychology classes from which the volunteers had been obtained. It was presented as a preliminary personality inventory. There was no indication that it was related in any way to the mental-set experiment.

Experiment II

Subjects.—Forty student volunteers from an introductory psychology class were tested individually in the set experiment. Thirty-six of these *Ss* received the anxiety scale in a group administration to the class from which they had previously been obtained. The class from which these *Ss* were obtained was not involved in Exp. I. The data to be presented were from these 36 *Ss*. The restricted *N* was due to absences from the class.

Procedure.—Five-letter anagrams typed in lower case on white 3 × 5-in. index cards were used as the problems. The anagrams employed may be found in an earlier article by Maltzman and Morrisett (**4**).

Three different classes of anagrams were used: eating-set, nature-set, and neutral anagrams. Most of the eating- and nature-set anagrams had only one possible solution, a word referring to some aspect of nature or eating; for example, "tulip" and "cakes." In the few cases where alternative solutions could be found, the alternative was a member of the same class of responses.

The *Ss* in the set phase of the experiment were randomly assigned to two different groups. Half of the *Ss*, designated the strong habit-strength group, received 15 eating- and 15 nature-set anagrams in a prearranged order. The remaining *Ss*, in the weak habit-strength group, received 26 neutral anagrams followed by two eating- and two nature-set anagrams in an ABBA order. These four anagrams were selected at random for each *S* from those used with the strong habit-strength group. A test series of ten anagrams followed immediately after the training series. Half of the *Ss* within each group were tested with eating-set anagrams while half received nature-set anagrams.

A 2-min. time limit was permitted for each anagram during the training and test series. If *S* did not attempt a solution within this interval or gave an incorrect solution, the correct solution was given by *E*. The instructions given to *Ss* have been described elsewhere (**4**).

Results

In Exp. I the anxiety scores obtained from the 48 *Ss* that received the water-jar problems ranged from 0 to 38. The median was 16.0, mean = 17.5, $SD = 8.5$. Using the same form of the anxiety scale Spence and Taylor (**6**) obtained scores ranging from 0 to 41. Since there was no obvious selective factor present owing to the failure to administer the anxiety scale to all experimental *Ss* in the two experiments reported here, the distributions obtained were probably not unduly biased.

A *tau* rank-order correlation coefficient (**10**) of −.27 was obtained be-

TABLE 1

EXPERIMENT I: ANXIETY SCORES AND DIRECT SOLUTIONS IN PROBLEMS 6–10

Condition	Anxiety				Direct Solutions			
	Mean	Median	*SD*	Range	Mean	Median	*SD*	Range
High anxiety	24.1	23.0	7.0	16–38	1.6	2.0	1.3	0–4
Low anxiety	11.0	12.0	4.3	0–16	2.7	3.0	1.9	0–5

tween the scores on the anxiety scale and the number of direct solutions on Problems 6–10. This is significantly different from zero at the .01 level of confidence. The 48 *Ss* that had taken the anxiety scale and had received the water-jar problem were divided into high- and low-anxiety drive groups. The 24 *Ss* with a median score or higher formed the high-anxiety group. The 24 *Ss* with a median score or lower formed the low-anxiety group. There were three scores at the median which were randomly assigned to the high and low conditions. The mean and median anxiety scores and direct solutions of the high- and low-anxiety groups on Problems 6–10 are presented in Table 1. A comparison of the number of direct solutions obtained on Problems 6–10 by *Ss* scoring in the upper and lower halves of the anxiety distribution gave a t of 2.42, $df = 46$, which is significant at the .02 level of confidence.

A chi-square test of independence was employed to determine whether Problem 8, the extinction problem that could be solved only by a direct solution, differentiated between *Ss* in the upper and lower halves of the anxiety distribution. Of the 24 *Ss* with high-anxiety scores 11 solved Problem 8 correctly with a direct solution. Of the 24 *Ss* with low-anxiety scores 17 solved Problem 8 correctly. A chi square of 2.14 corrected for continuity was obtained which is not significant ($p = .16$).

In Exp. II the anxiety scores obtained from the 36 *Ss* that received the anagram problems ranged from 4 to 34. The median was 16.5, mean = 16.2, $SD = 7.5$. There were 18 *Ss* in the high habit-strength group and an equal number in the low habit-strength group for which anxiety scores were available. The effects of the different initial habit strengths were reported in a previous paper (4).

Within each of the two habit-strength groups the *Ss* were divided into high- and low-anxiety groups. The *Ss* obtaining the nine lowest anxiety scores in each of the two habit-strength groups formed the low-anxiety groups and the *Ss* obtaining the nine highest scores formed the high-anxiety groups. Performance with the anagrams was scored in terms of

TABLE 2

EXPERIMENT II: ANXIETY SCORES AND ERRORS IN ANAGRAM TEST SERIES

Condition	Anxiety				Errors			
	Mean	Median	*SD*	Range	Mean	Median	*SD*	Range
High Anxiety	22.5	22.0	4.3	16–34	1.4	1.0	.84	0–3
Low Anxiety	9.8	9.5	3.6	4–17	2.7	2.5	1.7	0–5

errors on the ten anagram problems presented in the test series. Since the test anagrams had only one possible solution, any response other than the solution or the failure to give a response within 2 min. was scored as an error.

Since the error distributions were skewed and the variances were heterogeneous, a conventional factorial analysis of variance was not conducted. The x^2r test was employed instead (10). The high-anxiety groups made significantly fewer errors than the low-anxiety groups: $x^2r = 4.5$, $p = .04$. The effects of habit strength and the interaction between anxiety and habit strength were not significant. The mean anxiety scores and errors obtained by the high- and low-anxiety groups are shown in Table 2.

Discussion

The experiments reported here were concerned with some implications of Hull's assumption of a multiplicative relationship between habit strength and effective drive state for mental-set phenomena.

In Exp. I the dominant habit structure was acquired by responses congruent with indirect solutions. These responses were dominant in the hierarchy of alternative-solution responses, since they were the correct, reinforced, solutions for the first five problems. The indirect solutions would therefore tend to occur in Problems 6–10 where alternative direct solutions were possible or necessary for a correct solution. It is assumed here that an increase in the irrelevant need for anxiety resulted in a heightened drive state multiplying the habit strengths of the different classes of problem-solving responses. Assuming that the growth of habit strength is independent of the strength of the effective drive state (7), an increase in drive state would produce a greater absolute increase in effective reaction potential for the dominant class of responses, leading to an indirect or an incorrect solution. The significant negative correlation and t ratio obtained indicate that the greater the irrelevant anxiety need as indicated by the score on the scale employed, the less frequent the occurrence of an alternative direct solution. The difference in errors between the high- and low-anxiety groups on the extinction problem was not significant although in the predicated direction. The absence of a significant difference here may have been due to the failure to obtain a great enough differentiation in anxiety between the high- and low-anxiety groups.

In Exp. II *Ss* were trained with an equal number of nature and eating anagrams employing correction for errors. It is therefore assumed that the habit strength of these two classes of response solutions was approximately equal for each *S* at the conclusion of training. During the test series only one class of anagrams was presented to an *S*. Each anagram had only one possible solution. The occurrence of a correct solution for a given anagram would now be a function of the difference in strength of the reaction potential for the eating and nature classes of responses only one of which was correct. Other things being equal the difference between the reaction potentials of the correct and incorrect responses would be greater in the high-anxiety condition than in the low-anxiety condition. This is a consequence of the greater effective drive state in the former condition which would multiply the habit strengths for the correct and incorrect classes of responses. The significant difference obtained between the two

anxiety groups supports this implication of behavior theory.

Thus within the limits of the present studies an increase in manifest anxiety may produce an increase in indirect solutions and errors (Exp. I) or a decrease in errors (Exp. II) depending upon whether or not the dominant habit structure (mental set) is congruent with the class of correct solutions to a given class of problems.

The results of the present experiments are in accord with the verbal learning studies of Lucas (**2**) and Taylor and Spence (**9**). These studies were concerned with the influence of anxiety upon individual verbal responses while the present studies investigated the effects of anxiety upon classes of verbal responses elicited by classes of stimuli. The present studies further indicate that drive may influence the direction of thinking in accord with implications of Hull's behavior theory.

Summary

Some implications of Hull's assumption that irrelevant needs contribute to the effective drive state multiplying the habit strengths present were investigated. Two different problem-solving situations involving the establishment of mental sets were employed for this purpose. The water-jar problem was used in the first study to establish a set for a particular method of solution. It was found that the tendency to shift to a more direct method of solution was inversely related to the level of anxiety as determined by the scale employed.

In the second study anagrams having solutions referring either to some aspect of nature or eating were used to establish mental sets. During training *Ss* received an equal number of trials with eating- and nature-set anagrams. They were then presented with only one of these categories in the test series. *Ss* with high anxiety committed significantly fewer errors than *Ss* with low anxiety.

References

1. Cowen, E. L. The influence of varying degrees of psychological stress on problem-solving rigidity. *J. abnorm. soc. Psychol.*, 1952, **47**, 512–519.
2. Lucas, J. D. The interactive effects of anxiety, failure, and intra-serial duplication. *Amer. J. Psychol.*, 1952, **65**, 59–66.
3. Luchins, A. S. Mechanization in problem solving. *Psychol. Monogr.*, 1942, **54**, No. 6 (Whole No. 248).
4. Maltzman, I., & Morrisett, L. Different strengths of set in the solution of anagrams. *J. exp. Psychol.*, 1952, **44**, 242–246.
5. Maltzman, I., & Morrisett, L. The effects of task instructions on solution of different classes of anagrams. *J. exp. Psychol.*, 1953, **45**, 351–354.
6. Spence, K. W., & Taylor, J. A. Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.
7. Strassburger, R. C. Resistance to extinction of a conditioned operant as related to drive level at reinforcement. *J. exp. Psychol.*, 1950, **40**, 473–487.
8. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
9. Taylor, J. A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, **44**, 61–64.
10. Wilcoxon, R. *Some rapid approximate statistical procedures.* New York: American Cyanimid Co., 1949.
11. Yule, G. U., & Kendall, M. G. *An introduction to the theory of statistics.* New York: Hafner, 1950.

(Received October 25, 1952)

Journal of Experimental Psychology
Vol 46, No. 1, 1953

ANXIETY, ANXIETY REDUCTION, AND STRESS IN LEARNING[1]

JAMES DEESE, RICHARD S. LAZARUS, AND JAMES KEENAN

The Johns Hopkins University

This study is one in a series of experiments aimed at an analysis of the effects of stress upon performance. Previous studies have met with little success in analyzing individual differences in the effects of stress (2). In spite of the negative results obtained thus far, there is reason to believe that differences between individuals constitute the most important variable in the area of stress. We have also pointed out elsewhere (2) that some kind of interaction must exist between the type of stress and these individual differences.

Recently, a series of papers dealing with anxiety, learning, and avoidance conditioning suggested a possible lead in the study of individual differences and their interaction with type of stress. In the first of these studies, Taylor (6) selected two groups of *Ss* on the basis of a personality questionnaire which was presumed to measure "anxiety." She compared the rate of acquisition of a conditioned eyelid response in these two groups which she labeled "high anxiety" and "low anxiety," respectively. The "high anxiety" group conditioned much more rapidly than did the "low anxiety" group. A subsequent study (5) showed the same effect irrespective of the strength of the UCS. These data were interpreted exclusively in terms of the concept of anxiety as a drive, and anxiety reduction as reinforcement.

A more recent study from a different laboratory (1) reported a positive but nonsignificant correlation between essentially the same measure of "anxiety" and simple eyelid conditioning. The same study reported that *Ss* with "high anxiety" showed more stimulus generalization than did *Ss* with "low anxiety."

Somewhat different results have been found with more complicated material. Taylor and Spence (7) compared "high anxiety" and "low anxiety" with respect to performance on a kind of verbal maze. On this task, "low anxiety" was associated with rapid learning and fewer errors. This reversal was interpreted as being the result of the opportunity for interference allowed in the multiple-response situation. In other words, the authors postulated that the higher drive in the anxious *Ss* resulted in a decrement on this task because anxiety increases the strength of all responses, including competing ones.

In support of these findings is a paper by Montague (4) which shows that anxiety results in inferior learning when the task is such that there is great opportunity for interfering responses, but also leads to superior performance when such interfering responses are minimized. An experiment by Lucas (3) further confirms this interpretation.

The methods employed in the studies cited above seem applicable and useful in the study of the effects of psychological stress upon performance. Particularly significant is the fact that anxiety as a personality variable has been successfully used as a predictor of performance. The present study was designed to explore the relationship between anxiety as a personality variable and learning under various conditions of stress.

Procedure

The *Ss* were selected from a large group of students in introductory and social psychology at

[1] This research was supported in part by the United States Air Force under contract AF 33(038)-22624, monitored by the Perceptual and Motor Skills Research Laboratory of the Air Training Command Human Resources Research Center. Permission is granted for reproduction, translation, publication, use, and disposal in whole or in part by or for the United States Government.

The Johns Hopkins University. The selections were based on a personality questionnaire consisting of 30 items which were culled from the Minnesota Multiphasic Personality Inventory by Winne (8). These items had been chosen on the basis of the fact that they differentiated significantly between a neurotic and normal population, and the test was consequently called a "neuroticism inventory." Actually, this questionnaire was found to correlate highly with the "anxiety" questionnaire used by Taylor. A large number of items were identical.

The scores on the questionnaire varied from 0 to 17 in a distribution that was slightly skewed toward the high end. Approximately the upper and lower quartiles of this distribution were selected as the high and low groups, respectively. The scores for the low group ranged from 0 to 2, and those for the high group from 8 to 17.

The 90 *Ss* resulting from this division were divided into three experimental groups of 30 *Ss* each. In each experimental group there were 15 *Ss* scoring high on the inventory and 15 scoring low. All groups learned 12 consonant nonsense syllables by the method of anticipation for 12 trials.

The experimental conditions were as follows:

Control.—In this condition *Ss* were treated as they would be in a standard nonsense-syllable learning experiment, without threat or punishment of any kind. Each *S* was informed that he was a control for *Ss* in another experiment, and his cooperation was asked.

Avoidance-learning group.—In this condition, *Ss* were allowed to read through the list once, as in the case of the controls. But thereafter, they were administered a strong electric shock to the leg (produced by an inductorium) on the occurrence of incorrect responses or failures to respond. Thus, correct responses enabled them to avoid the electric shock. This fact was explained to each *S* beforehand.

Nonavoidance-learning group.—In this condition, the treatment of *Ss* was the same as in the previous condition except that the shocks were delivered irrespective of the correctness of *S*'s response. In other words, they were shocked as often when they were right as when they were wrong. Also, the frequency of the shocks was matched with the frequency for *Ss* in the avoidance-learning condition. The random nature of the shock was explained to *Ss* before the trials began.

TABLE 2

ANALYSIS OF VARIANCE OF NUMBER OF CORRECT RESPONSES†

Source	*df*	Mean Square	*F*
Within groups	84	200.41	
Conditions	2	260.01	1.30
Inventory	1	5397.55	26.90**
Inventory × Conditions	2	1978.71	9.87*
Total	89		

* Significant at .01 level.
** Significant at .001 level.
† Bartlett's test for homogeneity of variance yielded a chi square of 8.05, significant at approximately the .2 level.

RESULTS

Table 1 shows the mean total number of correct responses in all three conditions for both high and low *Ss*. Inspection of this table shows the high groups to give consistently more correct responses than the low groups. There is an interaction between the experimental conditions and the personality inventory such that the largest difference between the high and low groups occurs in the avoidance-learning situation.

Table 2 shows a statistical analysis of these data. The important thing to notice in this table is the highly significant difference between the learning scores of the high and low groups, and the highly significant interaction between the experimental

TABLE 1

LEARNING SCORES (CORRECT RESPONSES) FOR ALL GROUPS

Group	Control		Avoidance		Nonavoidance	
	M	*SD*	*M*	*SD*	*M*	*SD*
Anxious	31.40	15.52	37.07	20.88	27.53	19.67
Nonanxious	27.87	16.76	19.53	8.94	20.47	14.04

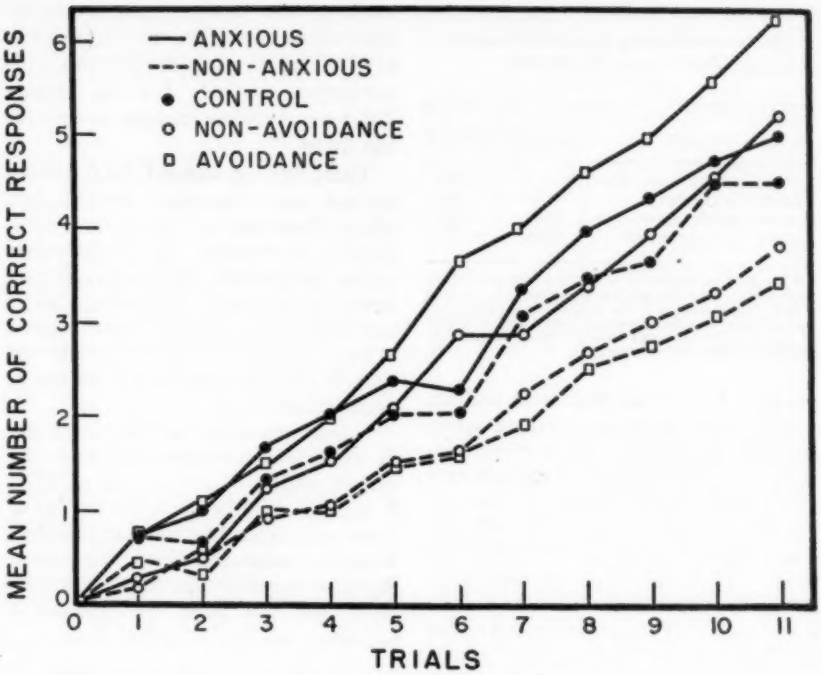

FIG. 1. Learning curves for all groups and conditions

conditions and the high and low groups.

The details of the learning curves are shown in Fig. 1. It is apparent from these curves that the difference between the high and low groups occurs early in learning and is maintained throughout the experiment. These curves also clearly show the nature of the interaction between conditions and the grouping based on the personality inventory. The differences between the high and low groups are strikingly large in the avoidance situation, somewhat smaller in the nonavoidance situation, and very small, though consistent, for the control condition. An inspection of these curves reveals one other important fact. The difference between the high and low groups is largely accounted for by an apparent decrement in performance on the part of the low group. This can be seen by looking at the bottom two curves in Fig. 1. (Also see Table 1.)

DISCUSSION

The primary finding in the present study, in partial confirmation of Taylor's original study, is that in avoidance learning there is a marked discrepancy in performance between "anxious" and "nonanxious" *Ss*. A major part of this difference cannot be accounted for in terms of anxiety as a drive (that is, in terms of improved performance of the "anxious" group), but rather must be interpreted in terms of a decrement in performance on the part of the "nonanxious"

TABLE 3

INTERCORRELATIONS BETWEEN VARIOUS PERSONALITY MEASURES (N = 227)

Variables	Including Overlap*	Excluding Overlap*
Anxiety vs. Psychasthenia	+.812	+.396
Anxiety vs. Hysteria	+.122	−.234
Hysteria vs. Psychasthenia	+.120	−.247
Anxiety vs. Neuroticism	+.686	+.554

* The first column refers to intercorrelations in which duplications of items are found among the various scales. Independence of the scales was accomplished by keeping the anxiety inventory intact and removing all other duplicated items from the hysteria, psychasthenia, and neuroticism scales.

group. In view of this latter finding, which could only be obtained when control conditions of performance were used, the concept of anxiety as a drive which Taylor brings to bear on her data must be partially rejected when applied to the present observations. Thus this concept loses considerable generality. It might also be mentioned that the conditioning curve for Taylor's (6) "nonanxious" group was very much more irregular than that found for the "anxious" group, an observation which is in keeping with the interpretation that the "low anxious" *Ss* suffered a decrement under stress.

Further light can be thrown on this issue if we compare in detail the learning curves for "high anxiety" and "low anxiety" groups in all of the three experimental conditions. Under the two conditions of shock (avoidance and nonavoidance), the "low anxiety" *Ss* fall below the comparable control groups in performance. For these two former conditions, the learning curves are almost identical (see Fig. 1). In the case of the "high anxiety" groups under the nonavoidance shock condition, performance is very much like the comparable control curves. In other words, there is no demonstrable effect of the stress upon the "high-anxiety" *Ss*. However, in the avoidance shock condition, the "high anxiety" *Ss* show some facilitation, thus accounting entirely for the greater difference between groups under this condition.

These findings suggest the following limited generalizations: In the first place, "nonanxious" *Ss* show a comparable decrement in performance under avoidance and nonavoidance stress conditions. Secondly, "anxious" *Ss* show some slight improvement in learning when the stress is such that that it can be avoided by means of performance.

It becomes apparent that all these data cannot be accounted for by the simple notion of anxiety as a drive. A number of other interpretations of these and other findings are possible. The most reasonable one depends upon the meaning of the personality inventory. On this point we thought it necessary to obtain further data.

After the above results were obtained, we retested most of the experimental and control *Ss* with the Taylor anxiety inventory. Moreover, 227 *Ss* from the introductory and social psychology classes at The Johns Hopkins University were administered the anxiety inventory along with the psychasthenia and hysteria scales from the MMPI, as well as the neuroticism questionnaire which had been used for the present experiment. The purpose of this latter testing was to obtain intercorrelations between the various personality tests in an effort to study the nature of the anxiety and neuroticism inventories.

With regard to the retesting of *Ss* used in the stress experiment with the anxiety inventory, we found essentially the same results as with the neuroticism inventory. In other words, the results found in the present experiment are applicable not only to the neuroticism inventory, but to Taylor's anxiety inventory as well. This latter result is reasonable in view of the correlation of +.686 between the neuroticism and anxiety inventories (see Table 3).

The intercorrelations between the various other inventories suggest further interpretation of our experimental data. Since all of the inventories included overlapping items, the intercorrelations were obtained both when these over-

lapping items were present and when they were deleted from each one of the scales. The Taylor anxiety inventory was always retained intact and the overlapping items were removed from each of the other inventories, respectively. Table 3 shows the results of this analysis.

As indicated by the kinds of items found in the anxiety and psychasthenia scales, clinical judgment considers them essentially alike. When overlapping items are included, the relationship is +.812, but when these items are removed, the *r* sinks to +.396. On the other hand, anxiety and hysteria, and hysteria and psychasthenia, are largely unrelated, depending somewhat upon the removal of overlapping items. We could be dealing here, in our breakdown of "high" and "low anxiety" groups, with a dimension of personality similar to Janet's hysteria-psychasthenia. Of course, this kind of analysis hardly exhausts the questions concerning the meaning of anxiety and the meaning of the other pencil-and-paper tests. At best it represents one possible beginning.

It is important to point out that there are other possible interpretations of the relationships that have been found between anxiety and learning and conditioning. It is hardly possible, at the present time, to choose definitively between these interpretations. However, the correlations we have obtained between anxiety scores and scores on the other personality tests suggest that any interpretation ought to be based upon an examination of the meaning of the inventories in terms of their intercorrelations.

For example, clinically, the hysterical individual is thought to respond to stress with great disturbance, while the psychasthenic is thought to approach emotional situations coolly and intellectually. On the basis of this relationship, we might predict that in a stressful situation the emotionally labile hysteric might show impairment in performance, while there would be little significant change in performance for the psychasthenic. This interpretation could account for the relative decrement of the nonanxious population in the present experiment. It does not exhaust the possibilities of interpretations in terms of personality variables and their dimensions.

In addition to the consideration of the meaning of the anxiety scores, it is important to consider the relationship between the present results and those found by Lucas (3) and Montague (4). In the latter cases, the effects of anxiety on performance were found to depend upon the nature of the learning task. The authors conclude that intraserial duplication or similarity led to poorer performance for "anxious" *Ss*. The lists in the present experiment consisted of consonant nonsense syllables of minimal similarity. The fact that different tasks and different populations,[2] seem to lead to different results means that task and population variables must be very thoroughly explored before the kinds of interpretations that have been offered can be considered generally applicable or of practical importance.

Summary

This study was designed to explore the relationships between certain problems of experimentally induced stress and personality factors. Ninety *Ss* were selected from a large population on the basis of extreme scores on a neuroticism inventory related to Taylor's (6) anxiety inventory. These *Ss* practiced a list of 12 consonant nonsense syllables under three different conditions: (*a*) shock administered for incorrect responses; (*b*) shock given at random for both correct and incorrect responses; (*c*) a control condition using no shock during learning.

In the avoidance situation (shock for incorrect responses), there was a large difference in the learning curves between the high-scoring and low-scoring groups in the "neuroticism" and "anxiety" inventories. A smaller difference between these two groups

[2] The present authors have obtained unpublished data indicating that with the identical design and task used in the present experiment, but with a population of lower intelligence level (and less homogeneous than the present college group in socioeconomic and educational status), high-anxious *Ss* perform more poorly than low-anxious *Ss*. This suggests that the "interference" interpretations (Lucas and Montague) of the reversals in performance of anxious and nonanxious *Ss* may be only one of a number of possible alternatives.

was found in the nonavoidance (random shock) condition. The control condition showed a very small though consistent difference between these conditions.

The large difference in the avoidance situation is due both to a decrement in the "low anxiety" group and a facilitation in the "high anxiety" group. In the nonavoidance situation, the difference is entirely due to a decrement in the "low anxiety" group.

These findings suggest that Taylor's interpretation of this kind of an effect in terms of an anxiety drive is too simple. An alternative hypothesis is suggested in terms of intercorrelations between the anxiety inventory and other pencil-and-paper tests.

References

1. Hilgard, E. R., Jones, L. V., & Kaplan, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, 1951, **42**, 94–100.
2. Lazarus, R. S., Deese, J., & Osler, S. F. The effects of psychological stress upon performance. *Psychol. Bull.*, 1952, **49**, 293–317.
3. Lucas, J. D. The interactive effects of anxiety, failure, and intraserial duplication. *Amer. J. Psychol.*, 1952, **65**, 59–66.
4. Montague, E. K. The role of anxiety in serial rote learning. *J. exp. Psychol.*, 1953, **45**, 91–96.
5. Spence, K. W., & Taylor, J. A. Anxiety and strength of UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–189.
6. Taylor, J. A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–93.
7. Taylor, J. A., & Spence, K. W. The relationship of anxiety level to performance in serial learning. *J. exp. Psychol.*, 1952, **45**, 61–64.
8. Winne, J. F. A scale of neuroticism: An adaptation of the Minnesota Multiphasic Personality Inventory. *J. clin. Psychol.*, 1951, **7**, 117–122.

(Received October 28, 1952)

Journal of Experimental Psychology
Vol 46, No. 1, 1953

NEGATIVE TRANSFER IN VERBAL LEARNING[1]

LYMAN W. PORTER[2] AND CARL P. DUNCAN

Northwestern University

It has usually been assumed that the optimum condition for producing negative transfer is that where *S* must learn successively two responses to the same stimulus (2). Thus with paired-associate verbal lists this situation is, in paradigm, A–B, A–C, by which is symbolized two successively learned lists with identical stimuli but different responses. Negative transfer reported with this design has not been great; usually the only evidence of interference in learning the second list is fewer correct responses in the first few trials on A–C as compared to a control condition.

Recently, Gagné, Baker, and Foster (1) have argued that with discriminative motor tasks the optimum condition for producing negative transfer is where the stimulus-response connections of the first task are reversed on the second task. The stimuli and responses of both tasks are identical; the second task is constructed by pairing each response with a stimulus different from the stimulus with which it was paired on the first task. We shall symbolize the two tasks in this case as A–B, A–B_R, where the subscript indicates reversal, by re-pairing, of the S–R connections. Although Gagné *et al.* were concerned with discriminative motor tasks, we shall assume that the same reasoning would apply to verbal lists, so we may speak of an A–B list and an A–B_R list.

It is the purpose of this paper to make a quantitative comparison of the amounts of negative transfer produced by the A–B, A–C and A–B, A–B_R methods.

[1] The advice of Dr. B. J. Underwood is gratefully acknowledged.

[2] Now at Yale University.

PROCEDURE

Subjects.—The *S*s were 40 Northwestern students all of whom had served in previous verbal-learning experiments. Practiced *S*s were used to reduce, in part, practice effects in learning successive lists, which would tend to mask negative transfer. Half of the group was assigned to the A–B, A–C condition (hereafter Group A–C), the other half to the A–B, A–B_R condition (Group A–B_R).

Method.—The design of the experiment is shown in Table 1. Group A–C learned, as their List II, entirely new responses to the same stimuli used in List I. Group A–B_R learned the same stimulus and response words on both lists, but in List II each response was paired with a stimulus different from the stimulus with which it had been paired on List I.

All *S*s used the same set of 12 two-syllable adjectives as the stimulus words of both lists. Two sets of two-syllable adjectives, 12 words in each set, were used as response words. With these three sets of words an A–B and a corresponding A–C list were constructed. Then the stimuli and responses of each of these lists were re-paired, resulting in an A–B_R list and an A–C_R list. Any one of these four lists may, of course, be used as List I, the first list learned, for either group. As shown in Table 1, each of the four lists was used equally often as List I in each group. The result is that each list was also used equally often as List II in each group. Thus any differences in difficulty of lists or differences

TABLE 1

DESIGN OF THE EXPERIMENT

Group	*S*s	List I	List II
A–B_R	5	A–B	A–B_R
	5	A–C	A–C_R
	5	A–B_R	A–B
	5	A–C_R	A–C
A–C	5	A–B	A–C_R
	5	A–C	A–B_R
	5	A–B_R	A–C
	5	A–C_R	A–B

TABLE 2

PERFORMANCE ON BOTH LISTS

Measure and List	Group A–B_R		Group A–C	
	M	σ_M	*M*	σ_M
List I				
Trials	13.95	1.33	13.65	1.09
Errors	13.35	2.64	13.25	2.08
List II				
Trials	17.85	1.57	12.80	1.77
Errors	22.15	3.34	7.10	1.07

due to stimulus-response pairing are controlled, and direct comparisons can be made both between groups of *S*s and from list to list within a group.

All lists were of low intralist similarity. They were presented on a Hull-type memory drum at a 2-sec. anticipation rate with 4 sec. between each trial. To prevent serial learning the pairs were presented in three different orders within a list.

All *S*s learned both lists to a criterion of one errorless trial. Two minutes elapsed between the learning of each list. Typical instructions for paired-associate learning were given before starting on List I. Before starting on List II all *S*s were told: "Now we are going to learn another list, the first words of which will be the same as the first words of the previous list. The second words will be different."

RESULTS

Table 2 shows measures of performance on both List I and List II for the two groups. It can be seen that the means of the groups for both trials and errors on List I differ by only fractional amounts. We can conclude, without testing for significance of differences, that the groups were well matched on List I.

Since we shall make comparisons between List I and List II within a single group, we need to know if there is any correlation between performance on the two lists. In Group A–B_R the correlation between the lists for trials was .44, for errors, .64. In Group A–C the correlation for trials was .67, for errors, .43. For 18 *df* the value of *r* needed for significance at the 5% level is .44; thus all the correlations reach or sufficiently approach significance.

Evidence for negative transfer can be gleaned from Table 2 by comparing performance on the two lists within a group. It is clear that there is no negative transfer in Group A–C; there is a decrease in both mean trials and mean errors from List I to List II. Since we are not concerned with positive transfer, we shall not test these decreases for significance.

In contrast, Group A–B_R shows presumptive evidence for negative transfer by an increase in both trials and errors on List II as compared with List I. The *t* for related measures was used to test these increases for significance. Since we are concerned only with differences in one direction; i.e., negative transfer, we shall consider all *t*'s as one-tailed tests. Comparison of the means for trials yielded $t = 2.53$, $p = .02$. Comparison of the error means yielded $t = 3.37$, $p < .01$. Thus the reversal group showed significant negative transfer by both measures.

Figure 1 shows learning of both lists for both groups by plotting successive criteria of correct responses as a func-

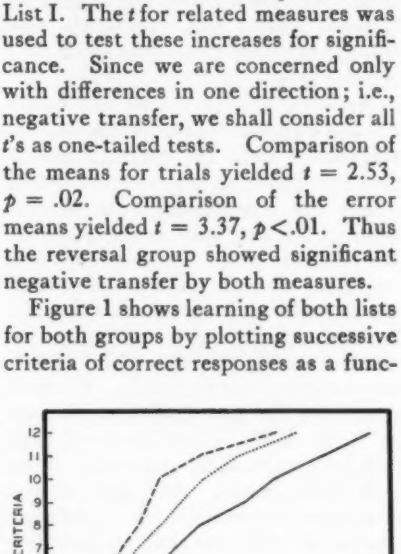

FIG. 1. Trials to reach successive criteria of correct responses on both lists

tion of mean trials to reach the successive criteria. The figure clearly indicates positive transfer to List II in Group A–C, and negative transfer in Group A–B_R.

Another index of transfer is the number of intrusions. As used here an intrusion is an error that consists of the response that was associated with a stimulus in List I being made to that stimulus during learning of List II. If that response was made (incorrectly) to any other stimulus in List II, it was called an error, but not an intrusion. During learning of List II, Group A–B_R made a total of 212 intrusions; Group A–C made 11. It is clear that the reversal group had great difficulty with intrusions of old responses.

Because of the large number of intruding responses made by the reversal group, their dat a were subjected to an item analysis to determine the relation between number of reinforcements (correct anticipations) on List I, and number and locus of intrusions on List II. The method was as follows. The total number of items was 240 (12 × 20 *S*s). These were divided into four reinforcement categories with the number of items in each category as nearly equal as possible. These categories were: 54 items with 1–4 reinforcements, 68 items with 5–7 reinforcements, 60 items with 8–11 reinforcements, and 58 items with 12–28 reinforcements. Then those items in each reinforcement category that appeared as intrusions in List II were recorded. Further, by dividing, as closely as possible, each *S*'s trials to learn List II into thirds, the locus of each intrusion was noted; i.e., whether it occurred in the first, second, or last third of the trials to learn the second list. For example, Table 3, which presents the item analysis, shows there were 18 intrusions in the first third of trials on List II of those items that had received 1–4 reinforcements on List I.

TABLE 3

INTRUSIONS AND INTRUSIONS PER ITEM IN LIST II BY GROUP A–B_R

Position in List II	Reinforcements on List I				
	1–4	5–7	8–11	12–28	Total
First third	18 .33	25 .37	26 .43	19 .33	88 .37
Second third	16 .29	17 .25	25 .42	25 .43	83 .35
Last third	7 .13	6 .09	14 .23	14 .24	41 .17
Total	41 .76	48 .71	65 1.08	58 1.00	212 .88

Since the number of items in each reinforcement category is not equal, Table 3 shows both the number of intrusions and, immediately below, the intrusions per item. These intrusion-per-item ratios reveal that, in general, the more frequently reinforced items appear more often as intrusions. As to locus, intrusions are made more frequently early in learning the second list, becoming less as practice continues.

DISCUSSION

Since we have no data bearing on the question of why there is such clear-cut negative transfer in learning the re-paired list, we shall merely offer a suggestion. It may be noted that *both* the stimuli and responses of the reversed list have acquired, during learning of the first list, associations which may be sources of interference. During practice on the re-paired list not only do the stimuli tend to produce incorrect responses but it may also happen that the responses tend to make *S* think of the old, and now incorrect, stimuli. If so, then both members of each pair of words in the

re-paired list, all 24 words, are potential sources of interference. This contrasts with the A–B, A–C situation where the C responses do not bear interfering associations, at least none recently and deliberately acquired.

We should also like to suggest that greater unlearning of an A–B list may be produced by subsequent learning of A–B_R than A–C. If so, investigators of retroactive and proactive inhibition may find the re-pairing of stimuli and responses to form a second list a useful technique for studying unlearning and response competition.

Summary

Twenty *Ss* in each of two groups learned two 12-item paired-associate adjective lists to one errorless trial. For one group (A–C) the lists had identical stimuli but different responses. For the other group (A–B_R) the stimuli and responses of the first list were re-paired to provide the second list. Examination of learning of the second list for evidence of negative transfer gave the following results:

1. There was no indication of negative transfer in the group that learned the A–B, A–C lists. The second list was learned in less trials and fewer errors than the first list.

2. The group that learned A–B, A–B_R showed negative transfer. Both trials and errors were significantly greater on the re-paired list than on the first list. In addition, this group made a far greater number of intrusions during learning of the second list than the A–C group.

References

1. Gagné, R. M., Baker, K. E., & Foster, H. On the relation between similarity and transfer of training in the learning of discriminative motor tasks. *Psychol. Rev.*, 1950, **57**, 67–79.
2. Underwood, B. J. *Experimental psychology.* New York: Appleton-Century-Crofts, 1949.

(Received November 12, 1952)

PSYCHOLOGICAL MONOGRAPH

#270

1945

ON PROBLEM SOLVING

By

KARL DUNCKER

$2.50

This popular monograph is #270 of the Psychological Monograph series. It has been reprinted so that it is again available.

Third Printing

AMERICAN PSYCHOLOGICAL ASSOCIATION

1333 Sixteenth Street N. W.

Washington 6, D. C.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The issues listed below are available for sale

Year	Volume	Available Numbers						Price per Number	Price per Volume
1916	1	1	2	3	4	5	6	$1.25	$7.00
1917	2	1	2	–	4	5	–	$1.25	$5.00
1920	3	–	2	3	4	5	6	$1.25	$6.25
1921	4	1	–	3	–	5	6	$1.25	$5.00
1922	5	–	–	3	4	5	–	$1.25	$3.75
1923	6	–	–	–	–	–	–	——	——
1924	7	1	2	–	–	5	–	$1.25	$3.75
1925	8	1	–	3	4	–	6	$1.25	$5.00
1926	9	–	2	3	–	5	6	$1.25	$5.00
1927	10	–	–	3	4	5	6	$1.25	$5.00
1928	11	1	2	3	4	5	6	$1.25	$7.00
1929	12	1	2	3	4	5	6	$1.25	$7.00
1930	13	1	2	3	4	5	6	$1.25	$7.00
1931	14	1	2	3	4	5	6	$1.25	$7.00
1932	15	1	2	3	4	5	6	$1.25	$7.00
1933	16	1	2	3	4	5	6	$1.25	$7.00
1934	17	1	2	3	4	5	6	$1.25	$7.00
1935	18	1	2	3	4	5	6	$1.25	$7.00
1936	19	1	2	3	4	5	6	$1.25	$7.00
1937	20	1	2	3	4	5	6	$1.25	$7.00
	21	1	2	3	4	5	6	$1.25	$7.00
1938	22	1	2	3	4	5	6	$1.25	$7.00
	23	1	2	3	4	5	6	$1.25	$7.00
1939	24	1	2	3	4	5	6	$1.25	$7.00
	25	1	2	3	4	5	6	$1.25	$7.00
1940	26	–	2	3	4	5	6	$1.25	$6.25
	27	1	2	3	–	5	–	$1.25	$5.00
1941	28	1	2	3	4	5	6	$1.25	$7.00
	29	1	2	3	4	5	6	$1.25	$7.00
1942	30	1	2	3	4	5	6	$1.25	$7.00
	31	1	2	3	4	5	6	$1.25	$7.00
1943	32	1	2	3	4	5	6	$1.25	$7.00
	33	1	2	3	4	5	6	$1.25	$7.00
1944	34	1	2	3	4	5	6	$1.25	$7.00
1945	35	1	2	3	4	5	6	$1.25	$7.00
1946	36	1	2	3	4	5	6	$1.25	$7.00
1947	37	1	2	3	4	5	6	$1.25	$7.00
1948	38	1	2	3	4	5	6	$1.25	$7.00
1949	39	1	2	3	4	5	6	$1.25	$7.00
1950	40	1	2	3	4	5	6	$1.25	$7.00
1951	41	1	2	3	4	5	6	$1.25	$7.00
	42	1	2	3	4	5	6	$1.25	$7.00
1952	43	1	2	3	4	5	6	$1.25	$7.00
	44	1	2	3	4	5	6	$1.25	$7.00
1953	45	By subscription, $15.00 (2 vols.)						$1.50	——
	46								

Information about the Journal of Experimental Psychology: as noted in the table, the journal was not published in 1918 and 1919; from 1937 to 1943 it was published at the rate of two volumes per year; from 1944 to 1950, in one volume per year. Beginning in 1951 there are two volumes a year.

Information about prices: when all numbers are available, the price per volume through 1952 is $7.00. Beginning in 1953, the price per volume is $7.50. For incomplete volumes, the price is $1.25 for each available number. For foreign postage, $.25 per volume should be added. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of $ 50.00 and over
20% on orders of $100.00 and over
30% on orders of $150.00 and over

Address orders and subscriptions to

AMERICAN PSYCHOLOGICAL ASSOCIATION

1333 Sixteenth Street N.W. Washington 6, D. C.